Universität Paderborn

Signal and System Theory Group

Master's Thesis

Spectrum sensing for Cognitive Radio

Author: Lucía Ruiz

Supervisor: Dr. David Ramírez

Paderborn, 2015

"It is paradoxical, yet true, to say, that the more we know, the more ignorant we become in the absolute sense, for it is only through enlightenment that we become conscious of our limitations. Precisely one of the most gratifying results of intellectual evolution is the continuous opening up of new and greater prospects."

Nicola Tesla.

Acknowledgements

First of all, I would like to acknowledge my supervisor Prof. David Ramirez, who has given me the opportunity to work with him, for the continuous support throughout my thesis, his patience and the good advice he has always given me.

To my family, the most important part of my life: mamá, papá, Miguel, Esther, Antonio and Ariadna. Without you and your support I would not be who I am.

To my friends, from the longlife ones, to the friends I have met during this last years and of course, all the people I have known during this unforgettable experience in Germany this year.

I am specially thankful to the people who has offered me all the help and support I needed in every moment during the last months.

Contents

List of Figures

List of Tables

Abbreviations

CR	**C**ognitive **R**adio
SNR	**S**ignal to **N**oise **R**atio
PDF	**P**robability **D**ensity **F**unction
PSD	**P**ower **S**pectral **D**ensity
NP	**N**eyman **P**earson
GLRT	**G**eneralized **L**ikelihood **R**atio **T**est
ML	**M**aximum **L**ikelihood
i.i.d	**i**ndependent **i**dentically **d**istributed
QoS	**Q**uality **o**f **S**ervice
CWGN	**C**omplex **W**hite **G**aussian **N**oise
MIMO	**M**ultiple **I**nput **M**ultiple **O**utput

Nomenclature

x	Scalar
$\mathbf{x}$	Column vector
$\mathbf{X}$	Matrix
$\mathbf{I}$	Identity Matrix
$\mathbf{0}$	Zero matrix or vector
Q	Complementary cumulative distribution function of the Gaussian
$Q_{\chi_v^2}$	Right-tail probability for a χ_v^2 rand
$\mathbf{R}$	Toepliz Matrix
$\hat{(\cdot)}$	Estime of a parameter
$\odot$	Hadamard product
$\det(\mathbf{X})$	Determinant of a matrix
$\mathrm{tr}(\mathbf{X})$	Trace of a matrix
$\mathcal{F}(\cdot)$	Fourier transform

Preface

Nowadays, the radio spectrum is a limited resource in wireless communications. This limitation is due to the fixed spectrum allocation principle, which allocates a spectrum band to a licensed user, even if it is not being used. Therefore, it has been necessary to introduce a new communications paradigm, which takes advantage of unused licensed spectrum. This known paradigm known as cognitive radio can improve the wireless usage sharing the spectrum between the licensed users and secondary users.

There are three different paradigms of cognitive radio: *underlay*, *overlay* and *interweave* cognitive radio. The first two paradigms allow concurrent communication between the primary and secondary users, whereas in the *interweave* cognitive radio, the secondary user transmits only when the primary is inactive. In the next chapter, a brief overview about these three paradigms will be presented, and finally we will focus on the last paradigm: *interweave* cognitive radio.

The main idea of interweave cognitive radio is the dynamic spectrum access, which allows secondary users to access unused resources in frequency, time and space. These secondary users search for opportunistically access of the spectrum when the primary users are inactive. Many factors (like SNR, interference, etc.) must be taken into account to maintain the quality of the service and allow the transmission of any user without causing degradation. In order to avoid the interferences, the probability of detection must be very high to be able to interrupt the secondary transmission whenever the primary users become active [1, 2].

In the next chapters, several sensing methods in different scenarios will be studied. To do that, we will present different detectors discussing their characteristics and performance. The detectors have to detect whether the received signal is only noise or there is a signal which has been transmitted from the user, i.e. the objective is to differentiate between two hypotheses, as a *binary hypothesis testing problem*. Common applications for this problem can be Radar, Sonar, Image processing, Biomedicine, etc. due to the fact that they all need to detect the presence/absence of a transmitted signal.

As introduction to the detection problem in this document, we will mainly differentiate between two types of detectors. Firstly, we will take into account a *prior* knowledge of the probability density functions (PDF) of each hypothesis. When the PDF of the signal is completely known, the Neyman-Pearson detector (NP) is the optimal detector [3]. Secondly, we will analyze the performance of the energy detector, which decides whether a signal has been transmitted or not, based on the received signal's energy. This detector is easy to implement, but has some limitations when the SNR is below a threshold. Thus, we have to study other detectors with good performance in the presence of some unknown parameters. In case the receiver does not have perfect knowledge of some of the parameters of the signal's PDF, it will be necessary to estimate them as good as possible. For these cases we will review the common approaches to the testing problem, going over the generalized likelihood ratio test (GLRT) and the Bayesian approach.

Finally a deep analysis of the GLRT in Cognitive Radio will be presented, due to its good performance. The communication system to analyze will be a system, in which the transmitted signal has rank P, i.e. where the transmitted signal has P independent transmitted streams.

The GLRT detector will be formulated for different scenarios, depending on the noise distribution in the receiving antennas. First, the GLRT will be derived for systems, where the noises are independent and identically distributed (i.i.d) is considered. Second, the noise with different and unknown variances (non i.i.d). The GLRT detector will be obtained for systems where only a given frequency channel has to be sensed (Chapter 3) or frequency selective channels (Chapter 4).

The numerical results will be presented in Chapter 5, where the GLRTs proposed will be compared under different scenarios.

In summary, in this thesis we will study the generalized likelihood ratio test as a solution to the spectrum sensing in multiple-input multiple-output environments. The performance will be analyzed by means of Monte Carlo simulations.

Chapter 1

Introduction to Cognitive Radio

Since the utilization of wireless transmission has immensely increased over the last two decades, the available spectrum has been reduced due to the fixed allocation of the spectrum [4]. The radio spectrum is a resource assigned by governments, i.e. the spectrum is divided into distinct frequency bands and each is assigned to a specific communication use [5]. These licenses can completely manage their own spectrum to get the best quality-of-service (QoS) and avoid the interference in their services as good as possible. Some of those licensed frequency bands are the radio and television bands, cellular and satellite bands and the traffic air control bands.

An example to show the band assignments is the frequency allocation chart of the United States (see Figure 1.1)[1] , in which most frequency bands have been assigned and there is only few available frequency bands for new transmission services or products. Thus, the idea of cognitive radio has appeared as a solution to the scarcity of spectrum. Through this paradigm the spectrum utilization can be improved, by allowing the sharing of the frequency band between licensed (primary) and non licensed (secondary) users - also called noncognitive and cognitive users, respectively[2]. For that, it is necessary to use the information of the activity, channel conditions, codebooks or messages of the users with whom spectrum is shared. Advanced radio and complex signal processing technology will be used to support this shared communication, in order to maintain the performance and quality of service. As a result, the spectrum efficiency will be improved and the bandwidth will be enough to support the demand for higher data rates in the future.

In this introductory chapter, three different types of cognitive radio will be studied [4]. These three types are *underlay*, *overlay* and *interweave* cognitive radio. *Underlay* cognitive

[1] 'http://www.ntia.doc.gov/files/ntia/publications/2003-allochrt.pdf'

[2] From now on, the licensed users will be named *primary users* and the users with dynamic access will be *cognitive users*

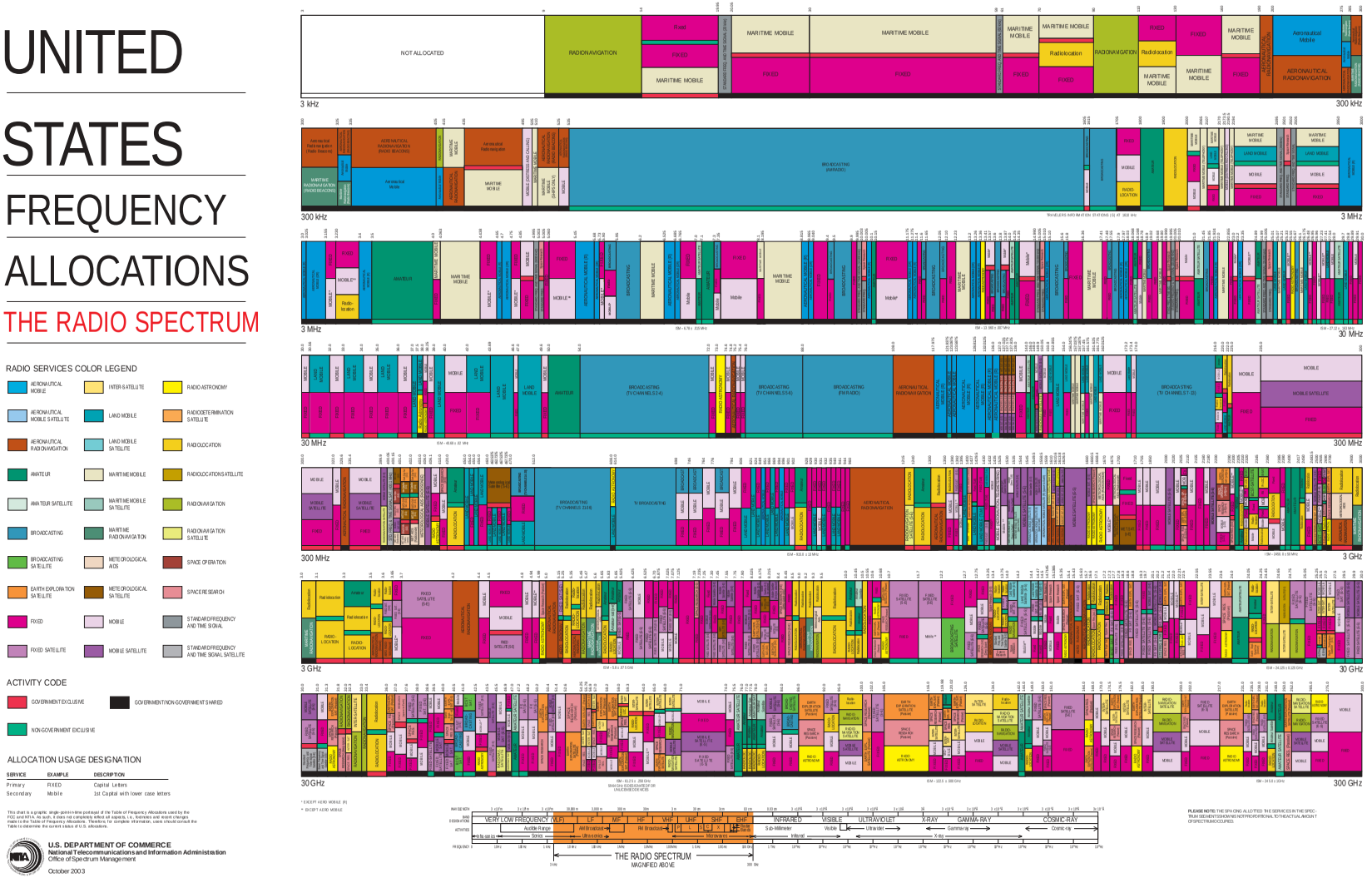

Figure 1.1: Frequencies allocation chart of the Unites States.

radio allows the secondary transmission only when the interference caused by the cognitive users does not exceed a fixed threshold. In *overlay* systems, the cognitive transmitter knows some information about primary users. This information is used to improve their communication while obtaining at the same time additional bandwidth for its own communication. Lastly the *interweave* cognitive radio transmits only when there are spectral holes in frequency and time, i.e. when the spectrum is eventually unused by the primary users.

A brief overview about these three different paradigms of cognitive radio will be presented in the introduction. We will finally focus on the analysis of *interweave cognitive radio*, which will be explained in depth in the following chapters, analyzing the results and commenting on them.

1.1 Underlay Cognitive Radio

In these systems, the interference caused to the primary users must be always known to the cognitive users to test if it always stays under an acceptable threshold. Only in that case will the transmission of the cognitive users be possible. To do this, the cognitive transmitter can approximate the interference caused in the primary user's signal, through overhearing the transmission from the cognitive receiver's location. Another option is that the cognitive user transmits with a limited power to maintain the interference as low as the threshold specifies.

Underlay communications with multiple cognitive and primary users are possible, and they can be formulated as a general multiuser communication problem with power constraints at the cognitive transmitters and interference constraints at primary receivers. In those systems, the power restriction will be limited by the sum of transmitting powers of the cognitive users. Thus, underlay systems are very restrictive whenever a peak interference power constraint at the primary users exist and as a result, the cognitive users must transmit below a certain power level.

1.2 Overlay Cognitive Radio

In overlay cognitive radio, the cognitive transmitter knows the channel information, the primary user's codebooks and sometimes also the messages. The primary user's message is obtained through the decoding at the cognitive's receiver. This knowledge of the primary user's signal can be used to cancel or mitigate the interference, i.e. the cognitive user assigns part of their power to improve the primary communications reducing the interference while transmitting their own information.

Different forms of cooperation are possible, in which the primary user works with different roles. The primary user might be *oblivious* to the cognitive user's presence and the decoding in the cognitive transmitter happens without assistance. On the other side, primary user can be *aware* of the cognitive user's presence and exploit the received signals from the cognitive users to cancel the interference. Moreover the primary user may be *partially or fully cooperative*, i.e. the primary user may sense the environment, forward messages and cooperate based on the obtained information. This last possibility obtains maximum benefits via cooperation.

1.3 Interweave Cognitive Radio

Interweave systems take advantage of the spectrum holes in time and frequency, i.e. whenever the spectrum is not occupied by the primary user. The spectrum efficiency will be improved making use of the opportunistic access to this non occupied bands. Since the communication is not concurrent, the primary users will not be interfered by the cognitive users. To make it possible, the radio spectrum must be periodically analyzed. This process is called spectrum sensing. To do that, it is necessary to make use of detectors and intelligent technology that allows the cognitive users to detect the presence/absence of the primary transmission in a spectrum band. The goal is to transmit whenever the primary user is inactive. The fading and shadowing effects are factors that make the detection a difficult task, due to the degradation of the signal.

Interweave cognitive radio will be from now on the topic of this thesis. We will study how spectrum sensing works and we will define different types of detectors and their characteristics.

In the Table 1.1, we will present the comparison of this three cognitive radio techniques, summarizing the characteristics previously explained.

Underlay	**Overlay**	**Interweave**
The channel information is known by the cognitive user. The caused interference to the primary user must be always known.	Cognitive nodes know the channel gains, codebooks and possibly the messages of the primary user. It may be used by the cognitive user to improve the primary communication.	Cognitive user has to sense the spectrum to seek spectral holes in space, time or frequency. Thus, the activity of the primary user must be monitored.
Concurrent transmission is possible when the interference caused by the secondary user is below a limit.	Simultaneous transmission is possible. The cognitive user can use part of its power to relay the primary user's message. The interference can be cancelled or mitigate by cooperation between both users.	The transmission is not concurrent and can be possible as long as the primary user is inactive. If the detector decides mistakenly that the primary user is idle, the transmission will be simultaneously. This probability of missed detection must be as low as possible.
Cognitive user's transmit power is limited by the interference constraint.	There is no restriction in the cognitive user's transmit power, since part of its power is used to relay the noncognitive user's message and therefore, improve the primary users's communication.	The power limitation is given by the range of cognitive user's spectral hole sensing.

Table 1.1: Comparison of underlay, overlay and interweave cognitive radio

Chapter 2

Spectrum Sensing for Cognitive Radio

Having explained the problem of spectrum bands scarcity and the solution given by cognitive radio in three different paradigms, we will focus in this chapter on one of those paradigms, which is the *interweave cognitive radio*. The goal is to transmit when the spectrum is unused by the primary user. To do that, it is necessary to detect the spectrum holes using spectrum sensing. In spectrum sensing, the secondary user has to detect periodically whether a primary transmission exists in a certain spectrum band and, whenever it is unused, the secondary transmission will be possible.

Many factors must be taken into account to get a successful transmission system and for this reason the spectrum sensing is not an easy to solve task. Firstly, the required SNR for detection may be very low in case the transmission power of the main signal is low. However the secondary user should detect the primary transmission and cease the use of the channel to avoid interference. Secondly, the time dispersion could make the coherent detection unreliable when it is unknown and the multipath fading could cause fluctuations in the signal power. Thirdly, there exists noise power uncertainty due to the change of the noise level in the wireless system [1].

In this thesis, several sensing methods will be proposed with different assumptions [3]. In some cases, the signal detection will be possible only when the signal's probability density function (PDF) $p(\mathbf{x}; \mathcal{H}_i)$ under each hypothesis is known. In such cases, the detectors are optimal, however this case is rarely realistic. For deterministic signals we will review the *matched filter*, and for random signals the *energy detector*, which has some SNR constraints. In realistic communication systems, the information about the transmitted signal's PDF is not completely known and therefore it is necessary to develop approaches to solve the spectrum sensing problem. In this part we will review the *Bayesian approach* and the

generalized likelihood ratio test (GLRT). This GLRT detector will be thoroughly analyzed in the last part of this thesis as a solution to the detection in systems where the transmitted signal has rank P, i.e. there are P streams transmitted at the same time. The reception will be also multiple, with L antennas at the spectrum monitor.

The noise properties will be also a point to consider. In multiple-input multiple-output (MIMO) systems, which are used to increase the channel capacity or improve the transmission reliability, we have to take into account whether the noise is identically distributed at each of the antennas.

2.1 Spectrum sensing

Spectrum sensing is the process of detecting if a certain spectral band is used or not, which allows a secondary transmission in this spectrum hole. Since we want to decide between two possible hypotheses, signal plus noise present versus noise only present, we consider a *binary testing problem*. Let us define $\mathcal{H}_1$ as the presence of transmitted signal in the spectrum band and $\mathcal{H}_0$ the absence of transmitted signal i.e. the availability of the band:

$$\begin{aligned} \mathcal{H}_0 &: x[n] = v[n], \qquad n = 0, 1, \ldots, N-1, \\ \mathcal{H}_1 &: x[n] = s[n] + v[n], \quad n = 0, 1, \ldots, N-1, \end{aligned} \tag{2.1}$$

where $s[n]$ represents a primary user's signal, $v[n]$ is noise and n represents time.

Different types of common detectors used in this process will be explained and analyzed. To do that, we have to analyze the received signal and apply the test decision $T(\mathbf{x})$ to decide if the signal has been generated under $\mathcal{H}_1$ or under $\mathcal{H}_0$, i.e.

$$T(\mathbf{x}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma. \tag{2.2}$$

Finally the performance of each one will be compared using the receiver operating characteristics (ROC) curve, which shows the *probability of false alarm* P_{FA} versus the *probability of detection* P_D. The P_{FA} is defined as the probability of deciding $\mathcal{H}_1$ when the transmission has been produced under $\mathcal{H}_0$ and the P_D is the probability of decide $\mathcal{H}_1$ when the transmission has been produced under $\mathcal{H}_1$, i.e.

$$\begin{aligned} P_{FA} &= \Pr(\mathcal{H}_1; \mathcal{H}_0), \\ P_D &= \Pr(\mathcal{H}_1; \mathcal{H}_1). \end{aligned} \tag{2.3}$$

This can be rewritten

$$\begin{aligned} P_{FA} &= \Pr(T(\mathbf{x}) > \gamma; \mathcal{H}_0), \\ P_D &= \Pr(T(\mathbf{x}) > \gamma; \mathcal{H}_1). \end{aligned} \tag{2.4}$$

A threshold has to be fixed to get the desired P_{FA} , what is also a very difficult task. Our goal will be to get the minimum false alarm probability obtaining the best probability of detection, but as we will see in the next sections, an increase in the probability of detection causes an increase in the probability of false alarm. Thus, we have to obtain a compromise between P_{FA} and P_D, which depends on the selection of the threshold.

2.2 A brief review of detection theory

2.2.1 Detectors with known PDF

The problem to be considered in this section is the detection of a signal with a known probability density function. We will study two types of detectors, the *matched filter* for deterministic signals and the *energy detector* for random signals. To do this, we assume the problem as a *binary hypothesis test*, which can be written as follows

$$\begin{aligned} \mathcal{H}_0 &: x[n] = v[n] && n = 0, 1, \ldots, N-1, \\ \mathcal{H}_1 &: x[n] = h[n]s[n] + v[n] && n = 0, 1, \ldots, N-1, \end{aligned} \tag{2.5}$$

where $v[n]$ is a white complex zero-mean Gaussian noise (CWGN) with variance σ^2, i.e. $v[n] \sim \mathcal{CN}(0, \sigma^2)$. In this section we will analyze two different distributions of the signal $s[n]$. Firstly, the signal is defined as a known deterministic signal, where the resulting detector will be the *matched filter*. Secondly, the signal is defined as a random Gaussian process with a known covariance. In such case the detector will be named energy detector. Since we want to represent a real wireless system, we have to take into account the effects of the propagation environment on a radio signal. We will have to simulate the fading of the signal according to a *Rayleigh distribution* $h[n]$ with normalized power. The *SNR* is defined as follows

$$SNR = \frac{P_s}{P_v} = \frac{1}{\sigma^2}, \tag{2.6}$$

where the P_s is the normalized signal power and P_v is the noise power, which is equivalent to the noise variance σ^2.

The goal of a detector is to maximize the probability of detection P_D subject to a constraint probability of false alarm P_{FA}. We will study in the next sections different ways to make this decision.

2.2.1.1 Matched filter

In this section, the signal $s[n]$ defined in the hypothesis test (2.5) is a known complex signal and the noise $v[n]$ has a complex Gaussian distribution with variance σ^2 and zero

mean, i.e. $v[n] \sim \mathcal{CN}(0, \sigma^2)$. Thus, the signal $x[n]$ under each hypotesis is defined as

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}), \\ \mathcal{H}_1 &: \mathbf{x} \sim \mathcal{CN}(\mathbf{s}, \sigma^2\mathbf{I}), \end{aligned} \tag{2.7}$$

where $\mathbf{x} = [\mathrm{x}[0]\mathrm{x}[1]\cdots\mathrm{x}[N-1]]^T$, $\mathbf{s} = [\mathrm{s}[0]\mathrm{s}[1]\cdots\mathrm{s}[N-1]]^T$ and $\mathbf{0}$ denotes a vector that contains N zeros. This detector applies the Neyman Pearson's criterion to make its decision. The matched filter decides $\mathcal{H}_0$ if the likelihood ratio $\mathscr{L}(\mathbf{x})$ exceeds a threshold γ and $\mathcal{H}_1$ in other case. The likelihood ratio is defined as

$$\mathscr{L}(\mathbf{x}) = \frac{p(\mathbf{x};\mathcal{H}_1)}{p(\mathbf{x};\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma. \tag{2.8}$$

The PDF under each hypothesis is defined as a complex normal distribution

$$\begin{aligned} p(\mathbf{x};\mathcal{H}_0) &= \left(\frac{1}{\pi^N\sigma^{2N}}\right)\exp\left[-\frac{1}{\sigma^2}\mathbf{x}^H\mathbf{x}\right], \\ p(\mathbf{x};\mathcal{H}_1) &= \left(\frac{1}{\pi^N\sigma^{2N}}\right)\exp\left[-\frac{1}{\sigma^2}(\mathbf{x}-\mathbf{s})^H(\mathbf{x}-\mathbf{s})\right], \end{aligned} \tag{2.9}$$

where H is the hermitian operator, given by the complex conjugate transpose.

Thus, the likelihood ratio is

$$\mathscr{L}(\mathbf{x}) = \exp\left[-\frac{1}{\sigma^2}\left((\mathbf{x}-\mathbf{s})^H(\mathbf{x}-\mathbf{s}) - \mathbf{x}^H\mathbf{x}\right)\right] \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma. \tag{2.10}$$

Taking logarithms

$$\begin{aligned} \ln\mathscr{L}(\mathbf{x}) &= -\frac{1}{\sigma^2}\left((\mathbf{x}-\mathbf{s})^H(\mathbf{x}-\mathbf{s}) - \mathbf{x}^H\mathbf{x}\right) \\ &= -\frac{1}{\sigma^2}\left[-\mathbf{x}^H\mathbf{s} - \mathbf{s}^H\mathbf{x} + \mathbf{s}^H\mathbf{s}\right] \\ &= \frac{2}{\sigma^2}\mathrm{Re}(\mathbf{s}^H\mathbf{x}) - \frac{1}{\sigma^2}\mathbf{s}^H\mathbf{s} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \ln\gamma. \end{aligned} \tag{2.11}$$

Since $\mathbf{s}$ is known, we decide following the next criterion

$$T(\mathbf{x}) = \mathrm{Re}(\mathbf{s}^H\mathbf{x}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma', \tag{2.12}$$

or equivalently

$$T(\mathbf{x}) = \mathrm{Re}\left(\sum_{n=0}^{N-1} x[n]\mathrm{s}^*[n]\right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma', \tag{2.13}$$

where $\gamma' = \frac{\sigma^2}{2}\left[\ln\gamma + \frac{1}{\sigma^2}\mathbf{s}^H\mathbf{s}\right]$ and * denotes the conjugate of the signal. Figure 2.1 shows how this matched filter works.

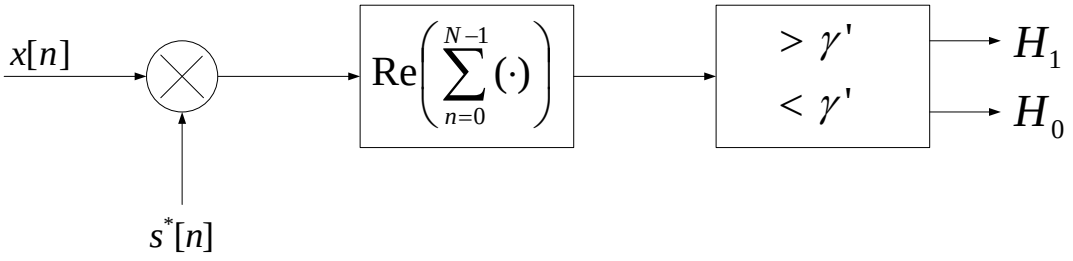

Figure 2.1: Matched filter for complex data

The performance is obtained by calculating the P_{FA} and P_D. To make it easier, we will let $z = \sum_{n=0}^{N-1} x[n]\mathrm{s}^*[n]$, which is a complex Gaussian random variable with different mean under each hypothesis

$$\begin{aligned} E(z;\mathcal{H}_0) &= \sum_{n=0}^{N-1} E(x[n])\mathrm{s}^*[n] = 0, \\ E(z;\mathcal{H}_1) &= \sum_{n=0}^{N-1} E(x[n])\mathrm{s}^*[n] = \sum_{n=0}^{N-1} |s[n]|^2. \end{aligned} \tag{2.14}$$

The variance of the signal z will be the same in both hypotheses because of the uncorrelation between the $x[n]$'s and taking into account that $\mathrm{var}(sz) = |s|^2\mathrm{var}(z)$. Then, the variance under $\mathcal{H}_0$ is

$$\begin{aligned} \mathrm{var}(z;\mathcal{H}_0) &= \mathrm{var}\left(\sum_{n=0}^{N-1} x[n]\mathrm{s}^*[n]\right) \\ &= \sum_{n=0}^{N-1} \mathrm{var}\left(x[n]\right)|s[n]|^2 \\ &= \sigma^2 \sum_{n=0}^{N-1} |s[n]|^2 = \sigma^2\mathcal{E}. \end{aligned} \tag{2.15}$$

The variance under $\mathcal{H}_1$ is calculated with the same procedure. The result will be the same as the previous one since the variance of $x[n]$ is the same under both hypotheses, as we can see in (2.7). This is

$$\mathrm{var}(z;\mathcal{H}_1) = \sigma^2\mathcal{E}. \tag{2.16}$$

As we can see, the distribution of the random variable z depends only on the signal energy and the noise variance. Thus, the hypothesis test $T(\mathbf{x})$ will be modeled as a Gaussian random variable whose mean will depend on the hypothesis and variance $\sigma^2\mathcal{E}$, this is

$$T(\mathbf{x}) = \mathrm{Re}(z) \sim \begin{cases} \mathcal{N}(0, \sigma^2\mathcal{E}/2) \text{ under } \mathcal{H}_0, \\ \mathcal{N}(\mathcal{E}, \sigma^2\mathcal{E}/2) \text{ under } \mathcal{H}_1. \end{cases} \tag{2.17}$$

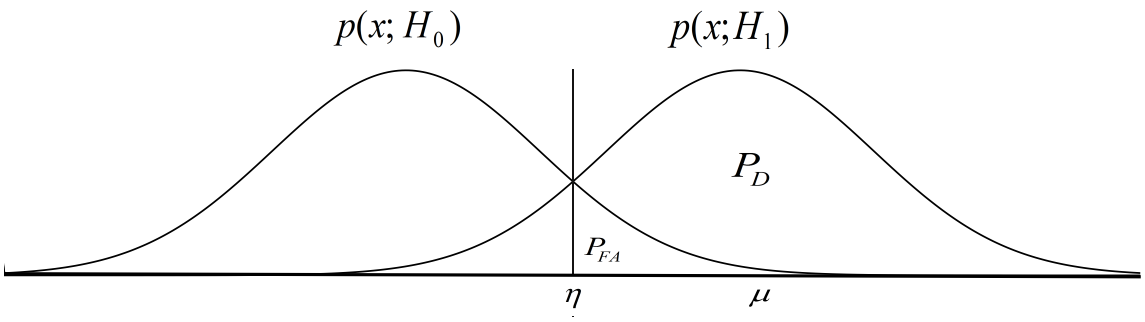

Figure 2.2: Hypothesis testing problem

With this distribution, we can calculate the P_{FA} and P_D, whose definition is given in (2.3), making use of the Gaussian distribution's properties. So, P_{FA} is defined in terms of Gaussian distribution, as the area of the $T(\mathbf{x})$'s PDF under $\mathcal{H}_0$ which is above the threshold and P_D is the area of the $T(\mathbf{x})$'s PDF under $\mathcal{H}_1$ which is above to the threshold, as we can see in the figure (2.2).

After some algebra, these probabilities are

$$\begin{aligned} P_{FA} &= \Pr(T(\mathbf{x};\mathcal{H}_0) > \gamma') \\ &= Q\left(\frac{\gamma'}{\sqrt{\sigma^2\mathcal{E}/2}}\right), \\ P_D &= \Pr(T(\mathbf{x};\mathcal{H}_0) < \gamma') \\ &= Q\left(\frac{\gamma' - \mathcal{E}}{\sqrt{\sigma^2\mathcal{E}/2}}\right), \end{aligned} \tag{2.18}$$

where Q is the complementary cumulative distribution function of the Gaussian, defined as follows

$$Q(\mathbf{x}) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}t^2 dt\right\}. \tag{2.19}$$

In figure 2.2 a generic hypothesis testing problem is represented, where η is the threshold and μ is the mean under $\mathcal{H}_1$.

2.2.1.2 Energy Detector

In the previous section the signal $\mathbf{s}$ was assumed deterministic and known. Such scenario is not common in real communication systems and therefore this detector is not a good solution to spectrum sensing, although the performance is optimal. In this section we will assume random signals and for that we can not assume the knowledge of the signal, but rather we have to assume that the signal is a random process with a known covariance structure. The studied detector for this assumption will be the energy detector, which

assume the same *binary hypothesis test* as (2.5) where $\mathbf{s}$ is defined as a Gaussian random process with zero mean and covariance matrix $\sigma_s^2\mathbf{I}$. The noise is complex white gaussian (CWGN) with covariance matrix $\sigma^2\mathbf{I}$. The signal $\mathbf{x}$ results

$$\begin{aligned}\mathcal{H}_0 &: \mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}),\\ \mathcal{H}_1 &: \mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \sigma_s^2\mathbf{I} + \sigma^2\mathbf{I}).\end{aligned} \tag{2.20}$$

The likelihood ratio test decides H_1 if

$$\mathscr{L}(\mathbf{x}) = \frac{p(\mathbf{x};\mathcal{H}_1)}{p(\mathbf{x};\mathcal{H}_0)} > \gamma, \tag{2.21}$$

where the PDFs are given by

$$\begin{aligned}p(\mathbf{x};\mathcal{H}_0) &= \frac{1}{pi^N \det\left(\sigma^2\mathbf{I}\right)} \exp\left[-\mathbf{x}^H \left(\sigma^2\mathbf{I}\right)^{-1}\mathbf{x}\right],\\ p(\mathbf{x};\mathcal{H}_1) &= \frac{1}{\pi^N \det\left(\sigma_s^2\mathbf{I} + \sigma^2\mathbf{I}\right)} \exp\left[-\mathbf{x}^H \left(\sigma_s^2\mathbf{I} + \sigma^2\mathbf{I}\right)^{-1}\mathbf{x}\right].\end{aligned} \tag{2.22}$$

This can be rewritten as

$$\begin{aligned}p(\mathbf{x};\mathcal{H}_0) &= \frac{1}{(\pi\sigma^2)^N} \exp\left[-\frac{1}{\sigma^2}\mathbf{x}^H\mathbf{x}\right],\\ p(\mathbf{x};\mathcal{H}_1) &= \frac{1}{(\pi(\sigma_s^2+\sigma^2))^N} \exp\left[-\frac{1}{(\sigma_s^2+\sigma^2)}\mathbf{x}^H\mathbf{x}\right],\end{aligned} \tag{2.23}$$

where the likelihood ratio test (2.21) results

$$\mathscr{L}(\mathbf{x}) = \frac{\frac{1}{(\pi(\sigma_s^2+\sigma^2))^N} \exp\left[-\frac{1}{(\sigma_s^2+\sigma^2)}\mathbf{x}^H\mathbf{x}\right]}{\frac{1}{(\pi\sigma^2)^N} \exp\left[-\frac{1}{\sigma^2}\mathbf{x}^H\mathbf{x}\right]} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma. \tag{2.24}$$

Taking logarithms, we have

$$\begin{aligned}\ln\mathscr{L}(\mathbf{x}) &= N\ln\left(\frac{\sigma^2}{\sigma_s^2+\sigma^2}\right) - \mathbf{x}^H\left[\frac{1}{\sigma_s^2+\sigma^2} - \frac{1}{\sigma^2}\right]\mathbf{x}\\ &= N\ln\left(\frac{\sigma^2}{\sigma_s^2+\sigma^2}\right) + \frac{\sigma_s^2}{(\sigma^2(\sigma_s^2+\sigma^2))}\mathbf{x}^H\mathbf{x} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \ln\gamma.\end{aligned} \tag{2.25}$$

Reducing this expression in data-dependent terms, the test statistic is

$$T(\mathbf{x}) = \mathbf{x}^H\mathbf{x} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma', \tag{2.26}$$

or, what is the same

$$T(\mathbf{x}) = \sum_{n=0}^{N-1} x^*[n]x[n] \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma', \tag{2.27}$$

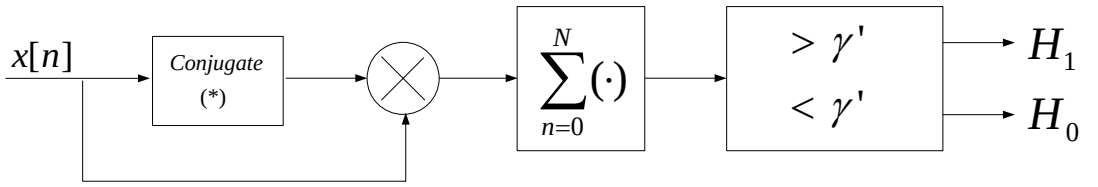

Figure 2.3: Energy detector for complex data

where

$$\gamma' = \frac{(\sigma^2(\sigma_s^2 + \sigma^2))}{\sigma_s^2}\left[\ln(\gamma) - N\ln\left(\frac{\sigma^2}{\sigma_s^2 + \sigma^2}\right)\right]. \tag{2.28}$$

Figure 2.3 shows the energy detector's diagram.

The performance of the detector is obtained by calculating the P_{FA} and P_D. To do that, let us define the distribution of $T(\mathbf{x})$ under each one of the hypotheses. Assuming $z = \sum_{n=0}^{N-1} \mathrm{x}^*[n]x[n]$, where the the mean is

$$\begin{aligned} E(z;\mathcal{H}_0) &= \sum_{n=0}^{N-1} E(x^*[n]x[n]) = 0, \\ E(z;\mathcal{H}_1) &= \sum_{n=0}^{N-1} E(x^*[n]x[n]) = 0. \end{aligned} \tag{2.29}$$

The variance under $\mathcal{H}_0$ and $\mathcal{H}_1$ is

$$\begin{aligned} \mathrm{var}(z;\mathcal{H}_0) &= \sigma^2, \\ \mathrm{var}(z;\mathcal{H}_0) &= \sigma_s^2 + \sigma^2. \end{aligned} \tag{2.30}$$

Thus, the distribution of the hypotheses test is defined as the sum of the squares of N *i.i.d* Gaussian random variables, which yields

$$\begin{aligned} \frac{T(\mathbf{x})}{\sigma^2} &\sim \chi_N^2 \text{ under } \mathcal{H}_0, \\ \frac{T(\mathbf{x})}{\sigma_s^2 + \sigma^2} &\sim \chi_N^2 \text{ under } \mathcal{H}_1. \end{aligned} \tag{2.31}$$

To define the P_{FA} and P_D we will make use of the right-tail probability $Q_{\chi_v^2}$ for a *chi-squared* χ_v^2 random variable distribution.

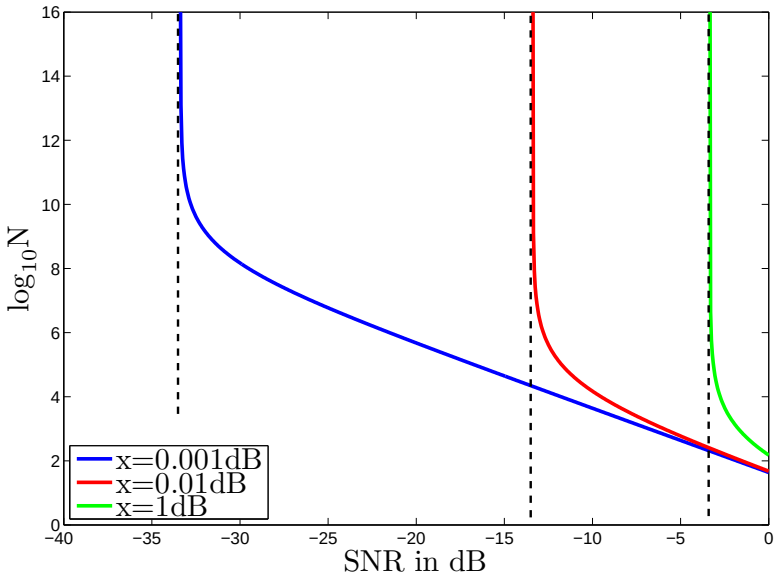

Figure 2.4: Number of samples N versus SNR, *SNR Walls*

Finally, according to (2.3) we have

$$\begin{aligned} P_{FA} &= \Pr(T(\mathbf{x};\mathcal{H}_0) > \gamma') \\ &= \Pr\left(\frac{T(\mathbf{x})}{\sigma^2} > \frac{\gamma'}{\sigma^2}\right) \\ &= Q_{\chi^2_N}\left(\frac{\gamma'}{\sigma^2}\right) \end{aligned} \tag{2.32}$$

and

$$\begin{aligned} P_D &= \Pr(T(\mathbf{x};\mathcal{H}_1) > \gamma') \\ &= \Pr\left(\frac{T(\mathbf{x})}{\sigma_s^2+\sigma^2} > \frac{\gamma'}{\sigma_s^2+\sigma^2}\right) \\ &= Q_{\chi^2_N}\left(\frac{\gamma'}{\sigma_s^2+\sigma^2}\right). \end{aligned} \tag{2.33}$$

This detector has some problems in the real-world, since we are assuming complete knowledge of the signal. In practice, some parameters are not known with an infinite precision. Specially the uncertainty in the noise variance might cause unrobustness in the detector[1]. The noise variance σ^2 is used to calculate the decision's threshold and therefore the P_{FA} is affected whenever there is some deviation in the value. In case the SNR is below to a certain value (*SNR Wall*), this deviation make the detection an impossible task to solve, even if the sensing duration is very high ($N \to \infty$) [6].

[1] A detector is nonrobust at a fixed SNR if it can not robustly achieve any pair (P_{FA},$P_{MD} = 1 - P_D$) where $0 < P_{FA} < 1/2$ and $0 < P_{MD} < 1/2$, even when N is arbitrary large.

Figure 2.4 shows how the N varies depending on the SNR value for a desired P_{FA} and P_D. As we can see, the value of N goes to infinite when the SNR is near the *SNR wall*. The variable x represents $x = 10\log_{10}\rho$, where $\rho > 1$ is a parameter that quantifies the noise uncertainty.

The energy detector limitations make us consider asynchronous detectors, which have to be robust even there exists noise uncertainty.

To summarize this subsection, the studied detectors with known PDF assumption are compared in the Table 2.1.

	Signals x	**Likelihood ratio**	**Conclusions**
Matched filter	$\mathbf{s}$ deterministic $\mathbf{x} \sim \begin{cases} \mathcal{H}_0 : \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}) \\ \mathcal{H}_1 : \mathcal{CN}(\mathbf{s}, \sigma^2\mathbf{I}) \end{cases}$	$\mathscr{L} = \mathrm{Re}(\mathbf{s}^H\mathbf{x})$	Optimal but not realistic. Necessary to know the transmitted signal $\mathbf{s}$. Synchronous with $\mathbf{s}$, not reliable in the low-SNR region.
Energy detector	$\mathbf{s}$ random variable $\mathbf{x} \sim \begin{cases} \mathcal{H}_0 : \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}) \\ \mathcal{H}_1 : \mathcal{CN}(\mathbf{0}, (\sigma_s^2 + \sigma^2)\mathbf{I}) \end{cases}$	$\mathscr{L} = \mathbf{x}^H\mathbf{x}$	Asynchronous with the transmitted signal. Noise variance σ^2 must be known to calculate the threshold. The deviation in the estimation of σ^2 implies a degradation. Under a SNR Wall the detection is not possible.

Table 2.1: Comparison of detectors with known PDF

2.2.2 Detectors with unknown PDF

In the previous section, we have assumed signals with completely known PDF for both hypotheses but this does not represent a real communication system in most of cases. Some parameters of the transmitted signal may be unknown. For example, the receptor may not always perfectly know the frequency of the transmitted signal and therefore we can not assume knowledge of the PDF . Moreover, the noise will be modeled as white Gaussian noise but with an unknown variance.

The approaches are, first of all, the *Bayesian approach*, which assigns prior PDFs to the unknown parameters considering the unknown parameters as realizations of random variables. The second approach will be the *generalized likelihood ratio test* (GLRT), which estimates the unknown parameters by the maximum likelihood estimates (MLEs).

2.2.2.1 Bayesian Approach

The Bayesian assigns prior PDFs to the unknown parameters $\boldsymbol{\theta}$ and so, the PDFs of the data can be calculated as

$$\begin{aligned} p(\mathbf{x};\mathcal{H}_0) &= \int p(\mathbf{x} \mid \boldsymbol{\theta_0};\mathcal{H}_0)\ p(\boldsymbol{\theta_0})d\boldsymbol{\theta_0}, \\ p(\mathbf{x};\mathcal{H}_1) &= \int p(\mathbf{x} \mid \boldsymbol{\theta_1};\mathcal{H}_1)\ p(\boldsymbol{\theta_1})d\boldsymbol{\theta_1}, \end{aligned} \tag{2.34}$$

where $p(\mathbf{x} \mid \boldsymbol{\theta_i};\mathcal{H}_i)$ is the PDF of $\mathbf{x}$, conditioned on $\boldsymbol{\theta_i}$, assuming that $\mathcal{H}_i$ is true. Thus, the Bayesian approach test is

$$\mathscr{L}(\mathbf{x}) = \frac{\int p(\mathbf{x} \mid \boldsymbol{\theta_1};\mathcal{H}_1)\ p(\boldsymbol{\theta_1})d\boldsymbol{\theta_1}}{\int p(\mathbf{x} \mid \boldsymbol{\theta_0};\mathcal{H}_0)\ p(\boldsymbol{\theta_0})d\boldsymbol{\theta_0}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma. \tag{2.35}$$

This detector requires complex multidimensional integration to solve the problem, which is not always possible in closed form. Therefore we will focus on the GLRT approach, which provides a good performance with a less complex solution.

2.2.2.2 Generalized Likelihood Ratio Test

The idea of this approach is to replace the unknown parameters by their maximum likelihood estimates (MLEs) $\hat{\boldsymbol{\theta}}$. The likelihood ratio test is

$$\mathscr{L}(\mathbf{x}) = \frac{p(\mathbf{x};\hat{\boldsymbol{\theta}}_{\mathbf{1}},\mathcal{H}_1)}{p(\mathbf{x};\hat{\boldsymbol{\theta}}_{\mathbf{0}},\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma, \tag{2.36}$$

where $\hat{\boldsymbol{\theta}}_0$ is the MLE of $\boldsymbol{\theta}_0$ assuming $\mathcal{H}_0$ and $\hat{\boldsymbol{\theta}}_1$ is the MLE of $\boldsymbol{\theta_1}$ assuming $\mathcal{H}_1$, i.e.

$$\begin{aligned} \hat{\boldsymbol{\theta}_{\mathbf{0}}} &= \underset{\boldsymbol{\theta_0}}{\operatorname{argmax}}\ p(\mathbf{x};\boldsymbol{\theta_0},\mathcal{H}_0), \\ \hat{\boldsymbol{\theta}_{\mathbf{1}}} &= \underset{\boldsymbol{\theta_1}}{\operatorname{argmax}}\ p(\mathbf{x};\boldsymbol{\theta_1},\mathcal{H}_1). \end{aligned} \tag{2.37}$$

This detector will be analyzed in the next chapters, where it will be derived under different assumptions.

Finally, a summary of detectors with unknown PDF assumption is presented in the Table 2.2.

	Likelihood ratio	**Procedure**	**Conclusions**
Bayesian approach	$\mathscr{L} = \frac{\int p(\mathbf{x} \mid \boldsymbol{\theta_1}; \mathcal{H}_1)\ p(\boldsymbol{\theta_1}) d\boldsymbol{\theta_1}}{\int p(\mathbf{x} \mid \boldsymbol{\theta_0}; \mathcal{H}_0)\ p(\boldsymbol{\theta_0}) d\boldsymbol{\theta_0}}$	Assign prior PDF to unknown parameters.	Complex multidimensional integrals to solve.
GLRT	$\mathscr{L} = \frac{p(\mathbf{x}; \hat{\boldsymbol{\theta}}_\mathbf{1}, \mathcal{H}_1)}{p(\mathbf{x}; \hat{\boldsymbol{\theta}}_\mathbf{0}, \mathcal{H}_0)}$	Estimation of the unknown parameters.	Asynchronous. Good performance. Easy to solve.

Table 2.2: Comparison of detectors with unknown PDF

Chapter 3

Multichannel GLRT detection of stationary white processes

In the previous chapter the problem of spectrum sensing has been introduced. Moreover, we have briefly explained some detectors classified into two categories, depending on the *prior* knowledge of the signal.

In the first category, the knowledge of the signal's probability density function is assumed to be complete. The optimal performance is provided by the matched filter, but it is necessary to know the transmitted signal and to be synchronized with the transmitter, that is not common in real communication systems. The second detector in this category is the energy detector, which uses the received signal's energy to make its decision. This detector is easy to implement and does not need synchronization with the transmitted signal. The main problem is caused by the noise variance uncertainty, which is required to calculate the threshold and therefore has a direct relation with the value of false alarm probability as we can see in (2.18). Thus, any deviation in the assumed noise variance with respect to the real value causes degradation in the detection performance. When the SNR is under a given SNR (*SNR Wall*), this detector will be not able to detect even if the sensing time is infinite.

Since the perfect knowledge of the signal's PDF is not common in spectrum sensing, in the second category we have studied some approaches to the testing problem. As we have shown, the Bayesian approach is difficult to implement because it requires complex multidimensional integration. The second approach is the generalized likelihood ratio test, which provides a less complex solution with a good performance by estimating the unknown parameters. Thus, in this thesis we will analyze the GLRT detector as a solution of spectrum sensing for cognitive radio [7].

In this chapter, we will review the performance of the GLRT detector in a multiple-input multiple-output (MIMO) communication system, with P transmitters and L receivers. The noise covariance is assumed unknown and has to be estimated by maximum likelihood estimation. To do that, the problem will be divided into different assumptions. Firstly, the GLRT detector will be analyzed in a scenario where the noise is independent and identically distributed (i.i.d) at each of the receivers. Secondly, the GLRT will be studied assuming non independent and identically distributed (non-i.i.d) noises.

3.1 Formulation problem

Let us consider a system where the transmitted signal has rank P. The spectrum monitor is equipped with L antennas, which has to sense the spectrum in a given frequency channel. The received signal $\mathbf{x} \in \mathbb{C}^L$ will be exploited to formulate the GLRT detector. The hypothesis testing problem is

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{x} = \mathbf{v}, \\ \mathcal{H}_1 &: \mathbf{x} = \mathbf{H}\mathbf{s} + \mathbf{v}. \end{aligned} \tag{3.1}$$

$\mathcal{H}_0$ is the hypothesis where there is no transmission, i.e. the received signal is only the noise $\mathbf{v} \in \mathbb{C}^L$. $\mathcal{H}_1$ is the hypothesis where a primary signal $\mathbf{s} \in \mathbb{C}^P$ has been transmitted and therefore, the received signal is the transmitted signal plus the additive noise. This noise is assumed as a zero-mean complex Gaussian spatially uncorrelated and temporally white. $\mathbf{H} \in \mathbb{C}^{L \times P}$ is the unknown multiple-input multiple-output (MIMO) channel between the primary user and the spectrum sensor. The signal $\mathbf{s}$ is assumed complex Gaussian temporally white and the power is normalized. We assume that the channel matrix $\mathbf{H}$ can absorb any spatial correlation and scaling of the primary signal. The covariance matrices of each signal are therefore

$$\begin{aligned} E[\mathbf{s}\mathbf{s}^H] &= \mathbf{I}_P, \\ E[\mathbf{v}\mathbf{v}^H] &= \mathbf{\Sigma}^2, \end{aligned} \tag{3.2}$$

where $\mathbf{I}_P$ is the identity matrix with size $P \times P$. The noise covariance matrix $\mathbf{\Sigma}^2$ is diagonal and unknown and σ_i^2 represents the noise variance in the $i-th$ component, i.e

$$\mathbf{\Sigma}^2 = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_L^2 \end{bmatrix}. \tag{3.3}$$

Since there are unknown parameters, the Neyman Pearson test can not be implemented and therefore, the GLRT approach will be used.

The detection problem is to decide between the hypotheses given in (3.1) testing the covariance matrix $\mathbf{R}$ of the received signal $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}_L, \mathbf{R})$. This matrix is defined under each of the hypotheses as follows

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{R}_0 = \mathbf{\Sigma^2}, \\ \mathcal{H}_1 &: \mathbf{R}_1 = \mathbf{H}\mathbf{H}^H + \mathbf{\Sigma^2}. \end{aligned} \tag{3.4}$$

Then, the GLRT based on the generalized likelihood ratio $\mathscr{L}$ is

$$\mathscr{L} = \frac{\max\limits_{\mathbf{R}_0} p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{R}_0)}{\max\limits_{\mathbf{R}_1} p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{R}_1)} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{3.5}$$

Having presented the GLRT test and the signals to be used, in the next sections the GLRT will be formulated under different scenarios.

3.2 Derivation of the GLRT for i.i.d. noises

In this first part, it is assumed a GLRT detection in which the noises at the received antennas are independent and identically distributed (i.i.d), i.e. the noises in (3.3) are $\sigma_i^2 = \sigma^2 \ \forall i$. Considering a sensing period, in which the channel remains constant and $M \geq L$ snapshots $\mathbf{x}_0, \cdots, \mathbf{x}_{M-1}$ are taken. It will be assumed as an independent identically distributed realizations of the signal $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}_L, \mathbf{R})$. The likelihood is calculated as the product of the individual PDFs

$$\begin{aligned} p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{R}) &= \prod_{i=0}^{M-1} p(\mathbf{x}_i; \mathbf{R}) \\ &= \frac{1}{\pi^{LM} \det(\mathbf{R})^M} \exp\left\{-M \mathrm{tr}\left(\hat{\mathbf{R}}\mathbf{R}^{-1}\right)\right\}, \end{aligned} \tag{3.6}$$

where $\hat{\mathbf{R}}$ is the sample covariance matrix:

$$\hat{\mathbf{R}} = \frac{1}{M} \sum_{m=0}^{M-1} \mathbf{x}_m \mathbf{x}_m^H. \tag{3.7}$$

In this particular case with i.i.d noises, the expression (3.5) can be rewritten with the values of $\mathbf{R}_i$ under each hypotheses, i.e. $\mathcal{H}_0 : \mathbf{R}_0 = \sigma^2\mathbf{I}$ and $\mathcal{H}_1 : \mathbf{R}_1 = \mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}$. Then, the generalized likelihood ratio can be written as

$$\mathscr{L} = \frac{\max\limits_{\sigma^2} p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \sigma^2)}{\max\limits_{\mathbf{H}, \sigma^2} p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{H}, \sigma^2)} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{3.8}$$

To obtain the closed form solution of this likelihood ratio, we have to estimate the value of the unknown parameters $\mathbf{R}_1$ and σ^2.

Firstly, we have to obtain the maximum likelihood (ML) estimate of σ^2 under $\mathcal{H}_0$, i.e. the value of σ^2 that makes the likelihood under the hypothesis $\mathcal{H}_0$ to be maximal. That is

$$\hat{\sigma}^2 = \underset{\sigma^2}{\operatorname{argmax}}\ p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \sigma^2). \tag{3.9}$$

To do that, we have to substitute $\mathbf{R} = \sigma^2\mathbf{I}$ in the likelihood (3.6)

$$\begin{aligned} p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \sigma^2\mathbf{I}) &= \frac{1}{\pi^{LM}\det(\sigma^2\mathbf{I})^M} \exp\left\{-M\mathrm{tr}\left((\sigma^2\mathbf{I})^{-1}\hat{\mathbf{R}}\right)\right\} \\ &= \frac{1}{\pi^{LM}\sigma^{2LM}} \exp\left\{-\frac{M}{\sigma^2}\mathrm{tr}\left(\hat{\mathbf{R}}\right)\right\}. \end{aligned} \tag{3.10}$$

and obtain the value of σ^2 that maximizes (3.10). Due to the monoticity of the logarithm function, the resulting optimization problem is

$$\begin{aligned} &\underset{\sigma^2}{\text{maximize}}\ -LM\log\pi - LM\log\sigma^2 - \frac{M}{\sigma^2}\mathrm{tr}\left(\hat{\mathbf{R}}\right), \\ &\text{subject to } \sigma^2 > 0, \end{aligned} \tag{3.11}$$

which can be solved with the Lagrange multipliers method [8]. To get the maximum value of σ^2, we have to apply the derivative of (3.11) and equate it to zero. The obtained value for $\hat{\sigma}^2$ is

$$\hat{\sigma}^2 = \frac{1}{L}\mathrm{tr}\left(\hat{\mathbf{R}}\right). \tag{3.12}$$

Thus, the expression of the likelihood under $\mathcal{H}_0$ can be written as

$$\log p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \hat{\mathbf{R}}_0) = -LM\log\pi - LM\log\left[\frac{1}{L}\mathrm{tr}\left(\hat{\mathbf{R}}\right)\right] - LM. \tag{3.13}$$

In order to estimate $\mathbf{R}$ under $\mathcal{H}_1$, we consider two cases depending on the rank P.

3.2.1 Sphericity Test: signals with rank $P \geq L - 1$

When the rank of the signal P is greater than or equal to the number of sensing antennas minus one, the covariance matrix $\mathbf{R}_1 = \mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}$ is positive definite Hermitian (i.e. $\lambda_i > 0\ \forall i$ and $\mathbf{R}_1 = \mathbf{R}_1^H$) and has no additional structure to exploit. Thus, the solution to the covariance matrix under $\mathcal{H}_1$ with this assumption is given by

$$\hat{\mathbf{R}}_1 = \hat{\mathbf{R}}, \tag{3.14}$$

where $\hat{\mathbf{R}}$ is the sample covariance matrix defined in (3.7). This result can be obtained by maximizing the likelihood under $\mathcal{H}_1$ with respect to $\mathbf{R}_1$, i.e.

$$\max_{\mathbf{R}_1} \log p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{R}_1), \tag{3.15}$$

where

$$\log p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{R}_1) = -LM \log \pi - LM \log \det(\mathbf{R}_1) - M \mathrm{tr}\left(\mathbf{R}_1^{-1} \hat{\mathbf{R}}\right). \tag{3.16}$$

To do this, we will apply the gradient to (3.16) with respect to $\mathbf{R}_1$ and equate the expression to $\mathbf{0}$ obtaining the estimate in (3.14). Thus, the likelihood under $\mathcal{H}_1$ is

$$\log p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \hat{\mathbf{R}}_1) = -LM \log \pi - LM \log \det(\hat{\mathbf{R}}) - ML. \tag{3.17}$$

Having obtained the estimates of the unknown values (3.12) and (3.14) we can substitute in (3.5) to obtain the solution of the likelihood ratio, which is a particular case of the GLRT approach called *Sphericity Test*

$$\log \mathscr{L} = ML \log \left[\frac{\det^{\frac{1}{L}}\left(\hat{\mathbf{R}}\right)}{\frac{1}{L}\mathrm{tr}\left(\hat{\mathbf{R}}\right)} \right] \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{3.18}$$

3.2.2 GLRT for signals with rank $P < L - 1$

When the rank $P < L - 1$, the matrix $\mathbf{R}_1$ has additional structure that allows us to improve the detection, due to the low rank structure of the primary signal. To estimate the value of $\mathbf{R}_1$ we will simplify the procedure to calculate the estimates, applying eigenvalue decompositions (EVDs).

Firstly, let $\mathbf{H}\mathbf{H}^H = \mathbf{U}\mathbf{\Psi}^2\mathbf{U}^H$ be an eigenvalue decomposition of $\mathbf{H}\mathbf{H}^H$, where $\mathbf{U} \in \mathbb{C}^{L \times P}$ is a unitary matrix which contains the eigenvectors, and

$$\mathbf{\Psi}^2 = \mathrm{diag}\left(\psi_1^2, \psi_2^2, \cdots, \psi_P^2, 0, 0, \cdots, 0\right), \tag{3.19}$$

with $\psi_1 \geq \psi_2 \geq \cdots \geq \psi_P$.

Secondly, to obtain the likelihood under $\mathcal{H}_1$, let us decompose with EVD the sample covariance matrix $\hat{\mathbf{R}} = \mathbf{W}\mathrm{diag}\left(\lambda_1, \cdots, \lambda_L\right)\mathbf{W}^H$ where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L$. With these

assumptions, the estimates of $\mathbf{U}$, $\mathbf{\Psi}^2$ and σ^2 are given by

$$\begin{aligned} \hat{\mathbf{U}} &= \mathbf{W}, \\ \psi^2 &= \lambda_i - \hat{\sigma}^2, \ i = 1 \cdots, P, \\ \hat{\sigma}^2 &= \frac{1}{L-P} \sum_{k=P+1}^{L} \lambda_k, \end{aligned} \tag{3.20}$$

whose demonstration is given in [9].

After some manipulations in (3.5) including (3.12) and using the previous EVD, the likelihood ratio detector is

$$\log \mathscr{L} = ML \log \left[\frac{\left(\prod_{i=1}^{L} \lambda_i \right)^{\frac{1}{L}}}{\frac{1}{L} \sum_{i=1}^{L} \lambda_i} \right] - M(L-P) \log \left[\frac{\left(\prod_{i=P+1}^{L} \lambda_i \right)^{\frac{1}{(L-P)}}}{\frac{1}{(L-P)} \sum_{i=P+1}^{L} \lambda_i} \right] \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma, \tag{3.21}$$

which is supposed to have the best performance in comparison with any other GLRT studied in this chapter, in case the variance of the noises is i.i.d. This is due to the additional structure of the covariance matrix and the possibility of exploiting the low rank structure of the primary signal. The first term of this expression is the statistic of the sphericity test (3.18) and the second one can be seen as the statistic of the sphericity test of the noise subspace.

3.3 Derivation of the GLRT fot non-i.i.d noises

In this section , the noise covariance matrix can not be simplified with a common variance as we did in the previous formulations. Thus, the parameters to estimate are $\mathbf{\Sigma}^2$ under $\mathcal{H}_0$ (defined in (3.3)) and the value of the covariance matrix $\mathbf{R}$ under $\mathcal{H}_1$. The particularization of (3.5) in this case is

$$\mathscr{L} = \frac{\max\limits_{\mathbf{\Sigma}^2} p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{\Sigma}^2)}{\max\limits_{\mathbf{R}_1} p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{R}_1)} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{3.22}$$

To estimate $\mathbf{\Sigma}^2$ under $\mathcal{H}_0$, we will apply the logarithm to $p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{\Sigma}^2)$. Taking into account the diagonal structure of $\mathbf{\Sigma}^2$, we will rewrite it as a product of the marginal PDFs where each $\left[\mathbf{\Sigma}^2\right]_{i,i}$ can be independently optimized. This is

$$\log p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \left[\mathbf{\Sigma}^2\right]_{i,i}) = -M \log \pi - M \log \left(\left[\mathbf{\Sigma}^2\right]_{i,i} \right) - M \frac{\left[\hat{\mathbf{R}}\right]_{i,i}}{\left[\mathbf{\Sigma}^2\right]_{i,i}}. \tag{3.23}$$

Applying the derivative in (3.23) and equating this expression to zero, we obtain the MLE of each independent (i,i) value as

$$\left[\hat{\mathbf{\Sigma}}^2\right]_{i,i} = \left[\hat{\mathbf{R}}\right]_{i,i}, \quad i = 1\ldots, L, \tag{3.24}$$

being $\left[\hat{\mathbf{R}}\right]_{i,i}$ the i,i element in $\hat{\mathbf{R}}$. Thus, the matrix $\hat{\mathbf{\Sigma}}^2$ is

$$\hat{\mathbf{\Sigma}}^2 = \hat{\mathbf{D}} = \operatorname{diag}\left(\left[\hat{\mathbf{R}}\right]_{1,1}, \ldots, \left[\hat{\mathbf{R}}\right]_{L,L}\right). \tag{3.25}$$

Finally the likelihood under $\mathcal{H}_0$ is

$$\log p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \hat{\mathbf{\Sigma}}^2) = -M \log \pi - M \log\det\left(\hat{\mathbf{D}}\right) - LM \tag{3.26}$$

To estimate the covariance matrix $\mathbf{R}$ under $\mathcal{H}_1$ the problem must be divided depending on the rank P as we did in the case of i.i.d noises.

3.3.1 GLRT for signals with rank $P \geq L - \sqrt{L}$

If the rank $P \geq L - \sqrt{L}$ [10], the covariance matrix $\mathbf{R}_1$ does not have additional structure to exploit, and therefore the ML estimate of $\mathbf{R}_1$ is given by the sample covariance matrix, as we saw in (3.14). The GLRT is now calculated making use of the expressions (3.17) and (3.26) as follows

$$\log \mathscr{L} = \log p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \hat{\mathbf{\Sigma}}^2) - \log p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \hat{\mathbf{R}}_1) = \log\left\{\frac{\det\left(\hat{\mathbf{R}}\right)}{\prod_{i=1}^{L}\left[\hat{\mathbf{D}}\right]_{i,i}}\right\}. \tag{3.27}$$

Since $\hat{\mathbf{D}}$ is defined by (3.25), the *Hadamard ratio* of the covariance matrix can be rewritten as follows

$$\mathscr{L} = \frac{\det\left(\hat{\mathbf{R}}\right)}{\prod_{i=1}^{L}\left[\hat{\mathbf{R}}\right]_{i,i}} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{3.28}$$

3.3.2 Signals with rank $P < L - \sqrt{L}$

If the signal has a rank $P < L - \sqrt{L}$ implies a low-rank structure in the primary signal, which can be exploited to improve the detection as we did in the Section 3.2.2. Let us

define the whitened sample covariance matrix and the whitened channel as

$$\begin{aligned}\hat{\mathbf{R}}_{\boldsymbol{\Sigma}} &= \boldsymbol{\Sigma}^{-1}\hat{\mathbf{R}}\boldsymbol{\Sigma}^{-1},\\ \mathbf{H}_{\boldsymbol{\Sigma}} &= \boldsymbol{\Sigma}^{-1}\mathbf{H},\end{aligned} \tag{3.29}$$

and rewriting the likelihood, we have

$$\begin{aligned}\log p(\mathbf{x}_0,\dots,\mathbf{x}_{M-1};\mathbf{H}_{\boldsymbol{\Sigma}},\boldsymbol{\Sigma}^2) = &- LM\log\pi - M\log\det\left(\mathbf{H}_{\boldsymbol{\Sigma}}\mathbf{H}_{\boldsymbol{\Sigma}}{}^H + \mathbf{I}\right)\\ &- M\log\det\left(\boldsymbol{\Sigma}^{\mathbf{2}}\right) - M\mathrm{tr}\left[\hat{\mathbf{R}}_{\boldsymbol{\Sigma}}\left(\mathbf{H}_{\boldsymbol{\Sigma}}\mathbf{H}_{\boldsymbol{\Sigma}}{}^H + \mathbf{I}\right)^{-1}\right].\end{aligned} \tag{3.30}$$

Now we have to obtain the MLEs of $\mathbf{H}_{\boldsymbol{\Sigma}}$ and $\boldsymbol{\Sigma}$ but due to the complexity, we will solve it making use of the eigenvalue decompositions of $\mathbf{H}_{\boldsymbol{\Sigma}}\mathbf{H}_{\boldsymbol{\Sigma}}{}^H$ and of the sample covariance matrix $\mathbf{R}_{\boldsymbol{\Sigma}}$.

First, the EVDs of $\mathbf{H}_{\boldsymbol{\Sigma}}\mathbf{H}_{\boldsymbol{\Sigma}}{}^H$ and $\hat{\mathbf{R}}_{\boldsymbol{\Sigma}}$ are

$$\begin{aligned}\mathbf{H}_{\boldsymbol{\Sigma}}\mathbf{H}_{\boldsymbol{\Sigma}}{}^H &= \mathbf{G}\boldsymbol{\Phi}^2\mathbf{G}^H,\\ \hat{\mathbf{R}}_{\boldsymbol{\Sigma}} &= \mathbf{Q}\mathrm{diag}\left(\gamma_1,\cdots,\gamma_L\right)\mathbf{Q}^H,\end{aligned} \tag{3.31}$$

where $\boldsymbol{\Phi}^2 = \mathrm{diag}\left(\phi_1,\cdots,\phi_L\right)$ and $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_L$. Then the ML estimates of $\mathbf{G}$ and $\boldsymbol{\Phi}^2$ are as follows

$$\begin{aligned}\hat{\mathbf{G}} &= \mathbf{Q},\\ \hat{\phi}_i^2 &= \begin{cases}\gamma_i - 1, & i = 1,\dots P,\\ 0 \quad\;\; , & i = P+1,\dots,L.\end{cases}\end{aligned} \tag{3.32}$$

The proof can be found in [9] as well as in the i.i.d case in the previous section (3.20). Making use of this obtained MLEs, the likelihood under $\mathcal{H}_0$ in (3.30) results

$$\begin{aligned}\log p(\mathbf{x}_0,\dots,\mathbf{x}_{M-1};\hat{\mathbf{H}}_{\boldsymbol{\Sigma}},\boldsymbol{\Sigma}^2) = &-LM\log\pi - M\sum_{i=1}^{P}\log\gamma_i\\ &- M\log\det\left(\boldsymbol{\Sigma}^2\right) - M\left(P + \sum_{i=P+1}^{L}\gamma_i\right)\end{aligned} \tag{3.33}$$

which can be rewritten as

$$\begin{aligned}\log p(\mathbf{x}_0,\dots,\mathbf{x}_{M-1};\hat{\mathbf{H}}_{\boldsymbol{\Sigma}},\boldsymbol{\Sigma}^2) = &- LM\log\pi - MP - M\log\det\left(\hat{\mathbf{R}}\right)\\ &- M\sum_{i=P+1}^{L}\left[\gamma_i - \log\gamma_i\right],\end{aligned} \tag{3.34}$$

The next step to obtain the GLRT is to maximize with respect to $\boldsymbol{\Sigma}^2$ in (3.33). If $P < L - \sqrt{L}$, there is not a closed-form solution. Thus, to calculate $\boldsymbol{\Sigma}^2$, we will study two

approaches: an alternating optimization method and a closed-form GLRT detector in the limit of asymptotically small SNR.

3.3.2.1 Alternating Optimization

The first proposed method to estimate when the rank $P < L - \sqrt{L}$ is an alternating optimization. The problem is partitioned into two different sets to obtain the alternatively minimization over each set of parameters maintaining the remaining ones with a fixed value. With this approach we can obtain a local minimum, since in each step the cost function always decreases, and therefore, converges. This method does not obtain the optimal value that maximizes the likelihood, but the result obtained by the detector has a good performance.

The ML optimization problem in (3.29) can be written as follows

$$\begin{aligned} &\underset{\mathbf{H_\Sigma},\mathbf{\Sigma}}{\text{minimize}}\ \text{tr}\left(\hat{\mathbf{R}}\mathbf{\Sigma}^{-1}\mathbf{R}_{\mathbf{\Sigma}}^{-1}\mathbf{\Sigma}^{-1}\right) - \log\det\left(\mathbf{\Sigma}^{-2}\right) + \log\det\mathbf{R}_{\mathbf{\Sigma}}, \\ &\text{subject to } \mathbf{R}_{\mathbf{\Sigma}} = \mathbf{I}_L + \mathbf{H}_{\mathbf{\Sigma}}\mathbf{H}_{\mathbf{\Sigma}}{}^H, \\ &\qquad\qquad [\mathbf{\Sigma}]_{i,i} \geq 0. \end{aligned} \tag{3.35}$$

To solve the problem with the alternating optimization, it has to be divided into two different sets, in which each one obtains the minimum value of $\mathbf{H_\Sigma}$ (fixed $\mathbf{\Sigma}$) and $\mathbf{\Sigma}$ (fixed $\mathbf{H_\Sigma}$) respectively, i.e.,

1) Minimization with respect to $\mathbf{H_\Sigma}$ *for fixed* $\mathbf{\Sigma}$.

The matrix $\hat{\mathbf{H}}_{\mathbf{\Sigma}}$ that minimizes the likelihood under $\mathcal{H}_1$ is

$$\hat{\mathbf{H}}_{\mathbf{\Sigma}} = [\mathbf{q}_1 \cdots \mathbf{q}_P]\left(\text{diag}\left(\gamma_1, \cdots, \gamma_P\right) - \mathbf{I}_P\right)^{\frac{1}{2}}, \tag{3.36}$$

where $\mathbf{q_i}$ is the *ith* column of $\mathbf{Q}$.

2) Minimization with respect to $\mathbf{\Sigma}$ *for fixed* $\mathbf{H_\Sigma}$

In this case, the optimization problem is

$$\begin{aligned} &\underset{\mathbf{\Sigma}}{\text{minimize}}\ \text{tr}\left(\hat{\mathbf{R}}\mathbf{\Sigma}^{-1}\mathbf{R}_{\mathbf{\Sigma}}^{-1}\mathbf{\Sigma}^{-1}\right) - \log\det\left(\mathbf{\Sigma}^{-2}\right), \\ &\text{subject to } [\mathbf{\Sigma}]_{i,i} \geq 0, \end{aligned} \tag{3.37}$$

which can be rewritten

$$\begin{aligned} &\underset{\boldsymbol{\alpha}}{\text{minimize}}\ \boldsymbol{\alpha}^T\left(\hat{\mathbf{R}}^T \odot \mathbf{R}_{\mathbf{\Sigma}}^{-1}\right)\boldsymbol{\alpha} - \sum_{i=1}^{L}\log\alpha_i^2, \\ &\text{subject to } \alpha_i^2 \geq 0, \end{aligned} \tag{3.38}$$

where $\boldsymbol{\alpha} = \left[\left[\boldsymbol{\Sigma}^{-1}\right]_{1,1}, \cdots, \left[\boldsymbol{\Sigma}^{-1}\right]_{L,L}\right]^T$ and the operator $\odot$ represents the Hadamard product, i.e. element to element. The matrix $\hat{\mathbf{R}}^T \odot \mathbf{R}_{\boldsymbol{\Sigma}}^{-1}$ is positive semidefinite and the problem is convex with respect to $\boldsymbol{\alpha}$, which makes it possible to solve it efficiently using a convex optimization solver. In the next chapter, where the performance results are presented, we will explain some possible convex optimization solvers to use in our simulations.

Having defined the two sets to minimize, the alternating optimization method has to be implemented following the next steps:

Alternating Optimization: Iterative Estimation of $\mathbf{H}_{\boldsymbol{\Sigma}}$ and $\boldsymbol{\Sigma}$

Input: Starting point $\boldsymbol{\alpha}_0$ and $\hat{\mathbf{R}}$

Output: ML estimates of $\mathbf{H}_{\boldsymbol{\Sigma}}$ and $\boldsymbol{\Sigma}$

Initialize: $n = 0$;

repeat

Compute $\boldsymbol{\Sigma}_{(n)}^{-1} = \text{diag}\left(\boldsymbol{\alpha}_{(n)}\right)$;

Obtain $\hat{\mathbf{R}}_{\boldsymbol{\Sigma}}^{(n+1)} = \boldsymbol{\Sigma}_{(n)}^{-1}\hat{\mathbf{R}}\boldsymbol{\Sigma}_{(n)}^{-1}$ and its EVD ;

Compute $\mathbf{H}_{\boldsymbol{\Sigma}}^{n+1}$ from (3.36) (fixed $\boldsymbol{\Sigma}_{(n)}^{-1}$) ;

Solve (3.38) to obtain $\boldsymbol{\alpha}_{(n+1)}$ (fixed $\mathbf{H}_{\boldsymbol{\Sigma}}^{(n+1)}$) ;

Update $n = n + 1$;

until convergence

The next step to solve the GLRT for non-i.i.d noises and $P < L - \sqrt{L}$ is to substitute the obtained estimates of $\boldsymbol{\Sigma}$ and $\mathbf{H}_{\boldsymbol{\Sigma}}$ in the likelihood under $\mathcal{H}_1$ given in (3.33) and apply the difference between the log-likelihoods under each hypothesis (3.26)-(3.33).

$$\log \mathscr{L} = \log p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \hat{\boldsymbol{\Sigma}}^2) - \log p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \hat{\mathbf{H}}_{\boldsymbol{\Sigma}}, \hat{\boldsymbol{\Sigma}}^2). \tag{3.39}$$

The complexity of this method might limit its utility. Thus, in the following section, we will study a simpler approach that assumes low SNR.

3.3.2.2 Low SNR approximation

This approach obtains a closed form expression for the GLRT in the low SNR regime. Such cases are very interesting in cognitive radio applications. The SNR will be close to zero, therefore the covariance matrix under $\mathcal{H}_1$ will be dominated by the noise covariance matrix, i.e. $\mathbf{R}_1 \approx \boldsymbol{\Sigma}^2$, where its MLE is $\hat{\boldsymbol{\Sigma}}^2 \approx \hat{\mathbf{D}}$, defined in (3.25). Thus, the likelihood

under $\mathcal{H}_1$ is

$$\begin{aligned} p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}, \hat{\mathbf{H}}, \hat{\boldsymbol{\Sigma}}^2) = &- LM \log \pi - MP - M \log \det \left(\hat{\mathbf{R}} \right) \\ &- M \sum_{i=P+1}^{L} \left[\beta_i - \log \beta_i \right], \end{aligned} \tag{3.40}$$

where the β_is are the eigenvalues of the sample spatial coherence matrix $\hat{\mathbf{C}} = \hat{\mathbf{D}}^{-1/2} \hat{\mathbf{R}} \hat{\mathbf{D}}^{-1/2}$, being $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_L$.

To get the log-likelihood ratio we have to calculate the difference between the likelihoods under $\mathcal{H}_0$ (3.26) and $\mathcal{H}_1$ (3.40). So, the log-GLRT closed form approach results as follows

$$\log \mathscr{L} \approx M \sum_{i=1}^{P} \left[\log \beta_i - \beta_i \right] + MP \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma, \tag{3.41}$$

or alternatively

$$\log \mathscr{L} \approx M \log \prod_{i=1}^{P} \beta_i e^{-\beta_i} + MP \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{3.42}$$

So far we have proposed different formulations of the GLRT problem in different scenarios depending on the noise distribution, where the sensing is applied only for a given frequency. The simulations and performance of each one will be analyzed in Chapter 6, comparing and commenting on the results with the help of the ROC curves.

Chapter 4

Multichannel GLRT detection of stationary processes with arbitrary PSDs

In the previous chapter the problem has been formulated as sensing in a given frequency. This problem will be extended to systems with frequency selective channels in which the signal's power spectral densities (PSDs) are unknown. The parameters to estimate in this section will be the PSDs $\mathbf{S}(e^{j\theta})$, instead of working with the covariance matrices $\mathbf{R}$ as we did previously.

In order to calculate the GLRT in the frequency domain, we will divide the problem into two different scenarios: equal PSDs and different PSDs, each one subdivided depending on the rank P of the signal. We will see in the next sections that the studied GLRT algorithms for frequency selective channels are the integral in frequency of the results obtained in (3.21) and (3.41) in the previous chapter.

4.1 Formulation problem

In environments with frequency-selective channels and arbitrary power spectral densities (PSD), the problem has to be reformulated as follows

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{x}[n] = \mathbf{v}[n], && n = 0, 1, \ldots, N-1, \\ \mathcal{H}_1 &: \mathbf{x}[n] = (\mathbf{H} * \mathbf{s})[n] + \mathbf{v}[n], && n = 0, 1, \ldots, N-1, \end{aligned} \tag{4.1}$$

where $\mathbf{s}[n] \in \mathbb{C}^P$ is the wide sense stationary (WSS) zero-mean complex Gaussian primary signal, $\mathbf{H}[n] \in \mathbb{C}^{L \times P}$ is the frequency selective multiple-input multiple-output (MIMO)

channel and $\mathbf{v}[n] \in \mathbb{C}^L$ is the WSS zero-mean circular complex Gaussian and spatially uncorrelated additive noise.. As in the previous chapter, any spatial and temporal correlation present in the signal can be absorbed in the unknown channel. The noise and signal covariance matrices are now defined as

$$\begin{aligned} &E[\mathbf{s}[n]\mathbf{s}^H[n-m]] : \mathbf{I}\delta[m], \\ &E[\mathbf{v}[n]\mathbf{v}^H[n-m]] : \mathbf{\Sigma}^2[m]. \end{aligned} \tag{4.2}$$

i.e. the signal covariance matrix function is an identity matrix with dimension $P \times P$ for $m = 0$ whereas the noise covariance matrix $\mathbf{\Sigma}^2[m]$ is a $L \times L$ diagonal matrix for all m, i.e.

$$\mathbf{\Sigma}^2[m] = \begin{bmatrix} \sigma_1^2[m] & 0 & \cdots & 0 \\ 0 & \sigma_2^2[m] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_L^2[m] \end{bmatrix}. \tag{4.3}$$

The data matrix $\mathbf{X}$ is defined as the combination of each $\mathbf{x}[n] = [x_1[n], x_2[n], \ldots, x_L[n]]^T$, where each $\mathbf{x}[n]$ vector is a column in the matrix $\mathbf{X}$. This is

$$\mathbf{X} = [\mathbf{x}[0]\ \mathbf{x}[1]\ \cdots\ \mathbf{x}[N-1]] \in \mathbb{C}^{L \times N}. \tag{4.4}$$

Therefore, the $i-th$ row contains N samples of the time series $x_i[n]$ at the $i-th$ antenna in the spectrum monitor. Let us define a vector $\mathbf{z}$ which contains the stacked columns of $\mathbf{X}$, i.e.

$$\mathbf{z} = \begin{bmatrix} \mathbf{x}[0] \\ \mathbf{x}[1] \\ \vdots \\ \mathbf{x}[N-1] \end{bmatrix} \in \mathbb{C}^{LN}. \tag{4.5}$$

Having defined the signal $\mathbf{z}$ necessary to solve the detection problem, the hypotheses can be written as

$$\begin{aligned} &\mathcal{H}_0 : \mathbf{z} \sim \mathcal{CN}\left(\mathbf{0}_{LN}, \mathcal{R}_0\right), \\ &\mathcal{H}_1 : \mathbf{z} \sim \mathcal{CN}\left(\mathbf{0}_{LN}, \mathcal{R}_1\right), \end{aligned} \tag{4.6}$$

whose block Toepliz covariance matrix $\mathcal{R} \in \mathbb{C}^{LN \times LN}$ has the form

$$\mathcal{R} = \begin{bmatrix} \mathbf{R}[0] & \mathbf{R}[-1] & \cdots & \mathbf{R}[-N+1] \\ \mathbf{R}[1] & \mathbf{R}[0] & \cdots & \mathbf{R}[-N+2] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}[N-1] & \mathbf{R}[N-2] & \cdots & \mathbf{R}[0] \end{bmatrix}. \tag{4.7}$$

and includes all the second order information of $\mathbf{x}[n]$, being each block of the covariance matrix defined as $\mathbf{R}[m] = E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right]$. As we saw in the last chapter, the covariance

matrices under each hypothesis are different, and are defined as follows

$$\begin{aligned}\mathcal{H}_0 &: \mathbf{R}_0[m] = \mathbf{\Sigma^2}[m], \\ \mathcal{H}_1 &: \mathbf{R}_1[m] = \sum_k \mathbf{H}[k]\mathbf{H}^H[k-m] + \mathbf{\Sigma^2}[m].\end{aligned} \tag{4.8}$$

4.2 Asymptotic Log-likelihood

To derive the log-GLRT for (4.8), it is necessary to obtain the ML estimates of the Toeplitz covariance matrices $\mathcal{R}$. Since it is a nonconvex problem, there is not a closed-form solution. Therefore, we propose the use of the *asymptotic likelihood* that is a function of the estimated and theoretical power spectral densities (PSDs). This asymptotic likelihood converges in the mean square sense to the time-domain likelihood, when the number of samples N goes to infinity, as Theorem 1 shows (the proof of this theorem can be found in [7]):

Theorem 1: As $N \to \infty$, the asymptotic log-likelihood converges in the mean square sense to the true log-likelihood.

$$\lim_{N\to\infty} E\left[\left|\frac{1}{N}\left[\log p(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathcal{R}) - \log p(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{S}(e^{j\theta}))\right]\right|^2\right] = 0. \tag{4.9}$$

The log-likelihood given in the first term is defined as a function of the block Toeplitz covariance matrix $\mathcal{R}$ (4.7)

$$\log p(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathcal{R}) = -NML\log\pi - M\log\det(\mathcal{R}) - M\mathrm{tr}\left(\hat{\mathbf{R}}\mathcal{R}^{-1}\right), \tag{4.10}$$

where the sample covariance matrix $\hat{\mathbf{R}}$ is given by:

$$\hat{\mathbf{R}} = \frac{1}{M}\sum_{i=0}^{M-1}\mathbf{z}_i\mathbf{z}_i^H. \tag{4.11}$$

The asymptotic log-likelihood in the second term is a function that depends on the power spectral density (PSD) matrix $\mathbf{S}(e^{j\theta}) = \mathcal{F}(\mathbf{R}[m])$, where $\mathcal{F}(\cdot)$ denotes a Fourier transform. Then, we have

$$\begin{aligned}\log p(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{S}(e^{j\theta})) = &-NML\log\pi - NM\int_{-\pi}^{\pi}\log\det\left(\mathbf{S}(e^{j\theta})\right)\frac{d\theta}{2\pi} \\ &- NM\int_{-\pi}^{\pi}\mathrm{tr}\left(\hat{\mathbf{S}}(e^{j\theta})\mathbf{S}^{-1}(e^{j\theta})\right)\frac{d\theta}{2\pi}.\end{aligned} \tag{4.12}$$

The variable $M(M > L)$ represents the number of independent realizations of the vector $\mathbf{z}$ produced in an experiment. The sample PSD $\mathbf{S}(e^{j\theta})$ is

$$\hat{\mathbf{S}}(e^{j\theta}) = \frac{1}{M}\sum_{i=0}^{M-1}\mathbf{x}_i(e^{j\theta})\mathbf{x}_i^H(e^{j\theta}), \tag{4.13}$$

where

$$\mathbf{x}_i(e^{j\theta}) = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\mathbf{x}_i[n]e^{-j\theta n}. \tag{4.14}$$

The hypothesis test defined in (4.6) can asymptotically be rewritten now as a hypothesis test for the power spectral density $\mathbf{S}(e^{j\theta})$ matrix

$$\begin{aligned}\mathcal{H}_0 &: \mathbf{x}(e^{j\theta}) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{S}_0(e^{j\theta})\right),\\ \mathcal{H}_1 &: \mathbf{x}(e^{j\theta}) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{S}_1(e^{j\theta})\right),\end{aligned} \tag{4.15}$$

being the PSDs under each hypotheses

$$\begin{aligned}\mathcal{H}_0 &: \mathbf{S}_0(e^{j\theta}) = \mathbf{\Sigma^2}(e^{j\theta}),\\ \mathcal{H}_1 &: \mathbf{S}_1(e^{j\theta}) = \mathbf{H}(e^{j\theta})\mathbf{H}^H(e^{j\theta}) + \mathbf{\Sigma^2}(e^{j\theta}),\end{aligned} \tag{4.16}$$

where $\mathbf{H}(e^{j\theta}) = \mathcal{F}\left(\mathbf{H}[n]\right)$ is the Fourier transform of the MIMO channel and $\mathbf{\Sigma^2}(e^{j\theta})$ is a diagonal matrix which contains the noises PSDs.

Then, the generalized likelihood ratio test $\mathscr{L}$ is asymptotically

$$\mathscr{L} = \frac{\max\limits_{\mathbf{S}_0(e^{j\theta})} p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{S}_0(e^{j\theta}))}{\max\limits_{\mathbf{S}_1(e^{j\theta})} p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{S}_1(e^{j\theta}))} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{4.17}$$

where, as in the previous chapter, it is necessary to find the ML estimates $\mathbf{S}_0(e^{j\theta})$ and $\mathbf{S}_1(e^{j\theta})$ that maximize the likelihoods.

This problem can be solved by finding the ML estimation of the matrix PSDs $\mathbf{S}(e^{j\theta})$ independently for each frequency, and finally applying the integral. Let us separate the problem depending on the noises distribution.

4.3 Derivation of the asymptotic GLRT for equal PSDs

In this section the noise PSDs are $\mathbf{S}_0(e^{j\theta}) = \mathbf{\Sigma^2}(e^{j\theta}) = S_v(e^{j\theta})\mathbf{I}$. Therefore, the likelihood ratio test in (4.17) is to find the ML estimations of $S_v(e^{j\theta})$ and $\mathbf{S}_1(e^{j\theta})$:

$$\mathscr{L} = \frac{\max\limits_{S_v(e^{j\theta})} p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; S_v(e^{j\theta}))}{\max\limits_{\mathbf{S}_1(e^{j\theta})} p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{S}_1(e^{j\theta}))} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \gamma. \tag{4.18}$$

To get the ML estimate of $\mathbf{S}(e^{j\theta})$ under $\mathcal{H}_0$, we have to calculate the value of $S_v(e^{j\theta})$ that maximizes the likelihood (4.12), that results in

$$\hat{S}_v(e^{j\theta}) = \frac{1}{L}\mathrm{tr}\left[\hat{\mathbf{S}}(e^{j\theta})\right]. \tag{4.19}$$

To get the ML estimation of $\mathbf{S}_1(e^{j\theta})$, the rank of the signal has to be taken into consideration. When the rank $P \geq L-1$, the problem does not have low rank structure. When $P < L-1$, there exists an additional structure that can be exploited to improve the detection. This reason implies the division of the problem in two parts.

4.3.1 GLRT for signals with rank $P \geq L-1$

When the rank is $P \geq L-1$, there is no low rank structure in the problem to exploit. Then, the value of $\mathbf{S}_1(e^{j\theta})$ that maximizes the likelihood under $\mathcal{H}_1$ is equal to (4.13), i.e.

$$\mathbf{S}_1(e^{j\theta}) = \hat{\mathbf{S}}(e^{j\theta}). \tag{4.20}$$

With this value, the asymptotic log-GLRT results as follows

$$\log\mathscr{L} = NML\int_{-\pi}^{\pi} \log\left[\frac{\left(\prod\limits_{i=1}^{L}\lambda_i(e^{j\theta})\right)^{\frac{1}{L}}}{\frac{1}{L}\sum\limits_{i=1}^{L}\lambda_i(e^{j\theta})}\right]\frac{d\theta}{2\pi} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \gamma, \tag{4.21}$$

which is the integral of the Sphericity test (3.18).

4.3.2 GLRT for signals with rank $P < L-1$

In case $P < L-1$ the ML estimation of $\mathbf{S}(e^{j\theta})$ is calculated following the same process as in the previous chapter, section (3.3.2). Then, to obtain the estimates under $\mathcal{H}_1$ we will

rewrite $\mathbf{H}(e^{j\theta})\mathbf{H}^H(e^{j\theta}) + S_v(e^{j\theta})\mathbf{I}$ as follows

$$\mathbf{H}(e^{j\theta})\mathbf{H}^H(e^{j\theta}) + S_v(e^{j\theta})\mathbf{I} = \mathbf{U}(e^{j\theta}) \left[\mathbf{\Psi}^2(e^{j\theta}) + \hat{S}_v(e^{j\theta})\mathbf{I}\right] \mathbf{U}^H(e^{j\theta}). \tag{4.22}$$

where $\mathbf{\Psi}^2(e^{j\theta}) = \text{diag}\left(\psi_1^2, \psi_2^2(e^{j\theta}), \cdots, \psi_P^2(e^{j\theta}), 0, 0, \cdots, 0\right)$ is a diagonal matrix of eigenvalues and $\mathbf{U}(e^{j\theta})$ the eigenvector matrix. Thus, the ML estimation of these variables are

$$\begin{aligned} \hat{\mathbf{U}}(e^{j\theta}) &= \mathbf{W}(e^{j\theta}), \\ \psi_i^2(e^{j\theta}) &= \lambda_i(e^{j\theta}) - \hat{S}_v(e^{j\theta}),\ i = 1 \cdots, P, \\ \hat{S}_v(e^{j\theta}) &= \frac{1}{L-P} \sum_{k=P+1}^{L} \lambda_k(e^{j\theta}). \end{aligned} \tag{4.23}$$

Where $\mathbf{W}(e^{j\theta})$ is a matrix that contains the eigenvectors of $\hat{\mathbf{S}}(e^{j\theta})$. The final expression of the *log-GLRT* is given by

$$\begin{aligned} \log \mathscr{L} =& NML \int_{-\pi}^{\pi} \log \left[\frac{\left(\prod_{i=1}^{L} \lambda_i(e^{j\theta}) \right)^{\frac{1}{L}}}{\frac{1}{L} \sum_{i=1}^{L} \lambda_i(e^{j\theta})} \right] \frac{d\theta}{2\pi} \\ & - NM\,(L-P) \int_{-\pi}^{\pi} \log \left[\frac{\left(\prod_{i=P+1}^{L} \lambda_i(e^{j\theta}) \right)^{\frac{1}{(L-P)}}}{\frac{1}{(L-P)} \sum_{i=P+1}^{L} \lambda_i(e^{j\theta})} \right] \frac{d\theta}{2\pi} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma, \end{aligned} \tag{4.24}$$

where $\lambda_i(e^{j\theta})$ is the $i-th$ largest eigenvalue of $\hat{\mathbf{S}}(e^{j\theta})$. We can see this log-GLRT (4.24) as the direct application of the log-GLRT (3.21) calculated in the previous chapter, integrating for each one of the frequencies.

4.4 Derivation of the asymptotic GLRT for different PSDs

In this case, the PSD under $\mathcal{H}_0$ is $\mathbf{S}_0(e^{j\theta}) = \mathbf{\Sigma^2}(e^{j\theta})$ and the GLRT is

$$\mathscr{L} = \frac{\max\limits_{\mathbf{\Sigma^2}(e^{j\theta})} p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{\Sigma^2}(e^{j\theta}))}{\max\limits_{\mathbf{S}_1(e^{j\theta})} p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{S}_1(e^{j\theta}))} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{4.25}$$

The ML estimate of the PSD under $\mathcal{H}_0$ is

$$\hat{\mathbf{\Sigma}}^2(e^{j\theta}) = \hat{\mathbf{D}}(e^{j\theta}) = \text{diag}\left(\left[\hat{\mathbf{S}}(e^{j\theta})\right]_{1,1}, \ldots, \left[\hat{\mathbf{S}}(e^{j\theta})\right]_{L,L} \right), \tag{4.26}$$

where $\hat{\mathbf{S}}(e^{j\theta})$ is defined in (4.13). Let us present the ML estimates of $\mathbf{S}(e^{j\theta})$ under $\mathcal{H}_1$ depending on the rank P to obtain the GLRT approaches.

4.4.0.1 GLRT for signals with rank $P \geq L - \sqrt{L}$

When the rank $P \geq L-\sqrt{L}$ there is no additional low rank structure. Then, the estimate of $\mathbf{S}_1(e^{j\theta})$ under $\mathcal{H}_1$ is the sample PSD $\hat{\mathbf{S}}(e^{j\theta})$ defined in (4.13). Then, the GLRT is obtained by introducing the estimates $\hat{\mathbf{\Sigma}}^2(e^{j\theta})$ (4.26) and $\hat{\mathbf{S}}_1(e^{j\theta})$ into the asymptotic likelihood ratio test (4.25), that results the integral of the Hadamard ratio (3.28), i.e.

$$\mathscr{L} = N \int_{-\pi}^{\pi} \frac{\det\left(\hat{\mathbf{S}}(e^{j\theta})\right)}{\prod_{i=1}^{L}\left[\hat{\mathbf{S}}(e^{j\theta})\right]_{i,i}} \frac{d\theta}{2\pi} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{4.27}$$

4.4.0.2 GLRT for signals with rank $P < L - \sqrt{L}$

In the case $P < L - \sqrt{L}$, there exists a low-rank structure to exploit and improve the detection. Due to the computational complexity of the alternating optimization in (3.3.2.1), we will only study the GLRT in the low SNR region, where we can find a closed form solution to the GLRT.

To estimate $\mathbf{S}_1(e^{j\theta})$ under $\mathcal{H}_1$, the first step is to define the whitened channel as $\mathbf{H}_{\mathbf{\Sigma}}(e^{j\theta}) = \mathbf{\Sigma}^{-1}(e^{j\theta})\mathbf{H}(e^{j\theta})$ and apply its eigenvalue decomposition as follows

$$\mathbf{H}_{\mathbf{\Sigma}}(e^{j\theta})\mathbf{H}_{\mathbf{\Sigma}}^{H}(e^{j\theta}) = \mathbf{G}(e^{j\theta})\mathbf{\Phi}^2(e^{j\theta})\mathbf{G}^{H}(e^{j\theta}), \tag{4.28}$$

with $\mathbf{\Phi}^2(e^{j\theta}) = \operatorname{diag}\left(\phi_1^2(e^{j\theta}), \phi_2^2(e^{j\theta}), \cdots, \phi_P^2(e^{j\theta}), 0, 0, \cdots, 0\right)$. Then the ML estimates of $\mathbf{G}(e^{j\theta})$ and $\phi_i^2(e^{j\theta})$ are given by

$$\begin{aligned} &\hat{\mathbf{G}}(e^{j\theta}) = \mathbf{Q}(e^{j\theta}), \\ &\hat{\phi}_i^2(e^{j\theta}) = \begin{cases} \beta_i(e^{j\theta}) - 1, & i = 1, \ldots, P, \\ 0 \qquad\qquad\quad , & i = P+1, \ldots, L. \end{cases} \end{aligned} \tag{4.29}$$

where $\beta_i(e^{j\theta})$ is the ith largest eigenvalue of $\hat{\mathbf{C}}(e^{j\theta}) = \hat{\mathbf{D}}^{-1/2}(e^{j\theta})\hat{\mathbf{S}}(e^{j\theta})\hat{\mathbf{D}}^{-1/2}$ and $\mathbf{Q}(e^{j\theta})$ is its eigenvector matrix. Finally the GLRT can be approximated as follows

$$\log \mathscr{L} \approx NM \int_{-\pi}^{\pi} \sum_{i=1}^{P} \log \beta_i(e^{j\theta}) \frac{d\theta}{2\pi} - NM \int_{-\pi}^{\pi} \sum_{i=1}^{P} \beta_i(e^{j\theta}) \frac{d\theta}{2\pi} + NMP \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{4.30}$$

That is the integral of the asymptotic log-GLRT given in (3.41).

Chapter 5

Simulations

In this chapter, we will evaluate the detection performance of the GLRT as a solution to the sensing problem in cognitive radio. To do that, we will simulate the scenarios proposed in Chapter 3 and Chapter 4 by means of Monte Carlo.

Firstly, we will summarize the likelihood ratio test of each detector, and the assumptions in its derivation. Secondly, we will represent some numerical results using the ROC curves and also evaluating the effect of the rank P in the performance. Moreover, in some cases we will represent the value of the missed detection probability $(1 - P_D)$ versus the value of the SNR.

5.1 Multichannel GLRT detection of stationary white processes

In frequency flat channels, the hypothesis test will be as follows:

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{x} = \mathbf{v}, \\ \mathcal{H}_1 &: \mathbf{x} = \mathbf{H}\mathbf{s} + \mathbf{v}. \end{aligned} \tag{5.1}$$

where the signal $\mathbf{s}$ has a complex white Gaussian distribution. The channel $\mathbf{H}$ is a multiple-input multiple-output (MIMO) channel of size $L \times P$ and the noise $\mathbf{s}$ is assumed complex white Gaussian. The SNR for this experiment is defined as follows

$$\text{SNR (dB)} = 10 \log_{10} \frac{\text{tr}(\mathbf{H}\mathbf{H}^H)}{\text{tr}(\mathbf{\Sigma}^2)}. \tag{5.2}$$

We will compare the algorithms derived under the assumptions given in Chapter 3. Firstly, GLRT detectors that assumes i.i.d noises. Secondly, GLRT detectors where the noise is assumed non-i.i.d.

Before we analyze the numerical results, a summary of the GLRT algorithms to compare is presented

GLRT for signals with i.i.d noises

- **Sphericity**

  GLRT derived under i.i.d noise assumption and rank of the transmitted signal $P \geq L-1$.

$$\log \mathscr{L} = ML \log \left[\frac{\det^{\frac{1}{L}} \left(\hat{\mathbf{R}} \right)}{\frac{1}{L} \mathrm{tr} \left(\hat{\mathbf{R}} \right)} \right] \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{5.3}$$

- **iid-GLRT**

  GLRT derived assuming i.i.d. noise and rank of the transmitted signal $P < L-1$.

$$\log \mathscr{L} = ML \log \left[\frac{\left(\prod_{i=1}^{L} \lambda_i \right)^{\frac{1}{L}}}{\frac{1}{L} \sum_{i=1}^{L} \lambda_i} \right] - M(L-P) \log \left[\frac{\left(\prod_{i=P+1}^{L} \lambda_i \right)^{\frac{1}{(L-P)}}}{\frac{1}{(L-P)} \sum_{i=P+1}^{L} \lambda_i} \right] \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{5.4}$$

- **iid-GLRT (P = 1)**

  GLRT algorithm assuming i.i.d noises and rank of the transmitted signal $P = 1$, whose likelihood ratio is

$$\log \mathscr{L} = ML \log \left[\frac{\mathrm{tr} \left(\hat{\mathbf{R}} \right)}{\lambda_1} \right] \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{5.5}$$

GLRT for signals with non-i.i.d noises

- **Hadamard**

  GLRT derived under non i.i.d noise assumption, where the rank $P \geq L - \sqrt{L}$

$$\mathscr{L} = \frac{\det \left(\hat{\mathbf{R}} \right)}{\prod_{i=1}^{L} \left[\hat{\mathbf{R}} \right]_{i,i}} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{5.6}$$

- **alternating**

  This algorithm does not have a closed form solution as the previous GLRTs. The ML estimates are found applying the alternating optimization algorithm given in **??**. Then, the likelihood ratio is

$$\log \mathscr{L} = \log p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \hat{\mathbf{\Sigma}}^2) - \log p(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \hat{\mathbf{H}}_{\mathbf{\Sigma}}, \hat{\mathbf{\Sigma}}^2). \tag{5.7}$$

- **asymp-GLRT**

  GLRT derived under non i.i.d noise assumption and rank of the transmitted signal $P < L - \sqrt{L}$. This approach is valid in the low SNR regime

$$\log \mathscr{L} \approx M \sum_{i=1}^{P} [\log \beta_i - \beta_i] + MP \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{5.8}$$

- **asym-GLRT (P=1)**

  This is a particular case of the previous algorithm, where only one signal is transmitted

$$\log \mathscr{L} \approx -\beta_1 \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{5.9}$$

5.1.1 Numerical results

5.1.1.1 Comparison of rank P versus missed detection probability - i.i.d noises

Figure (5.1) shows the missed detection probability $(1 - P_D)$ versus the rank P of the signal for i.i.d noises. The P_{FA} is fixed to $P_{FA} = 0.001$, the signal to noise ratio is SNR = -8dBs and the number of antennas is $L = 6$.

Since the *i.i.d GLRT* does not assume any structure on the primary signal, its performance will be the best one with respect the other detectors for arbitrary values of P.

Moreover, it is clear that the *i.i.d GLRT* and the *i.i.d GLRT ($P = 1$)* have the same performance when the rank is 1. When P increases, *i.i.d GLRT ($P = 1$)* suffers more degradation since we are assuming a false parameter P to derive the algorithm.

It should also be noted that when the rank is $P \geq L - 1$, the *Sphericity test* offers similar performance to the GLRTs. That is, taking into account that $L = 6$, the *Sphericity test* obtains its best performance when $P \geq 5$.

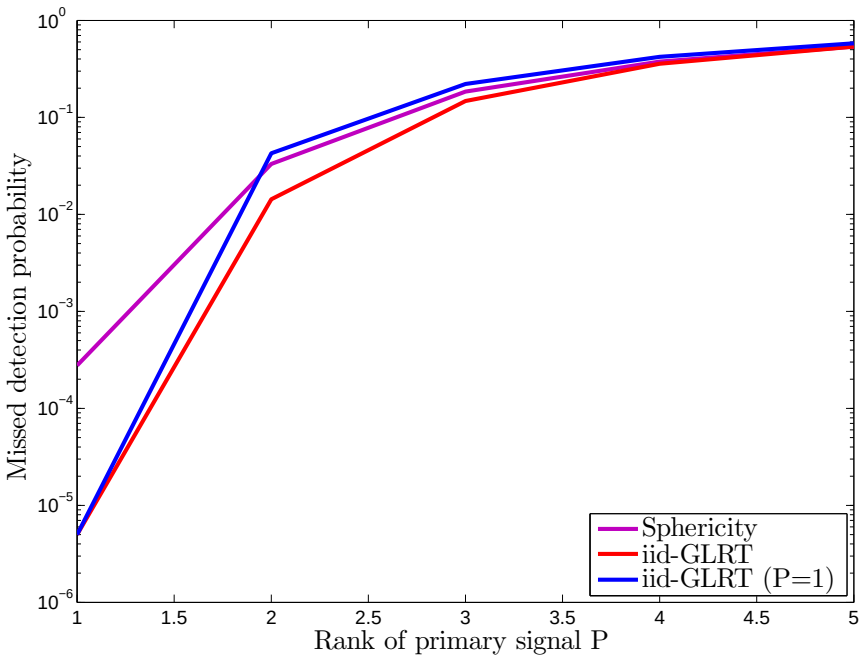

Figure 5.1: Missed detection probability versus P with $P_{FA} = 0.001$ and i.i.d noises

5.1.1.2 Comparison of rank P versus missed detection probability - non-i.i.d noises

The figure (5.2) presents a scenario whose parameters are the same as in the previous figure, i.e, SNR = - 8dB, a fixed $P_{FA} = 0.001$, and the number of antennas $L = 6$. Now the noise distribution is assumed non-i.i.d.

When the rank $P > 1$, we can see how the *asymp-GLRT* algorithm offers the the same detection performance as the *alternating* algorithm, which is computationally more complex since it needs to iterate until convergence. In case the rank $P \geq 4$, the *Hadamard* ratio test offers a similar performance to the *asymp-GLRT*

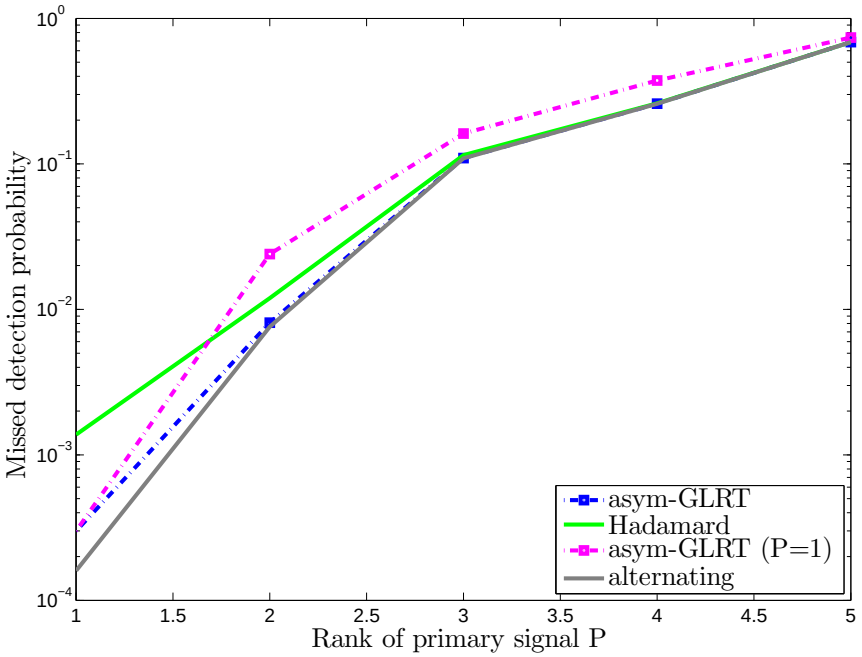

Figure 5.2: Missed detection probability versus P with $P_{FA} = 0.001$ and non i.i.d noises

5.1.1.3 Comparison of false alarm probability versus detection probability - i.i.d noises

Figure (5.3) compares all the GLRT algorithms derived in Chapter 3 using the ROC curves, which compare the detection probability versus the false alarm probability. In this case, the rank of the signal is $P = 1$, the SNR = -8dB, $L = 4$ antennas and it is assumed i.i.d noises.

As expected, the best results are given by the GLRT, whose design assumes i.i.d noises and exploits the low rank structure of the signal, i.e the *i.i.d-GLRT*. The *Sphericity* ratio test has also better performance as the rest of algorithms derived under non-i.i.d noise assumption.

It is interesting to note that the detectors designed for uncalibrated receivers, i.e *alternating*, *asym-GLRT* and *Hadamard* do not suffer a notable degradation in their performance, since they do not assume any noise covariance structure.

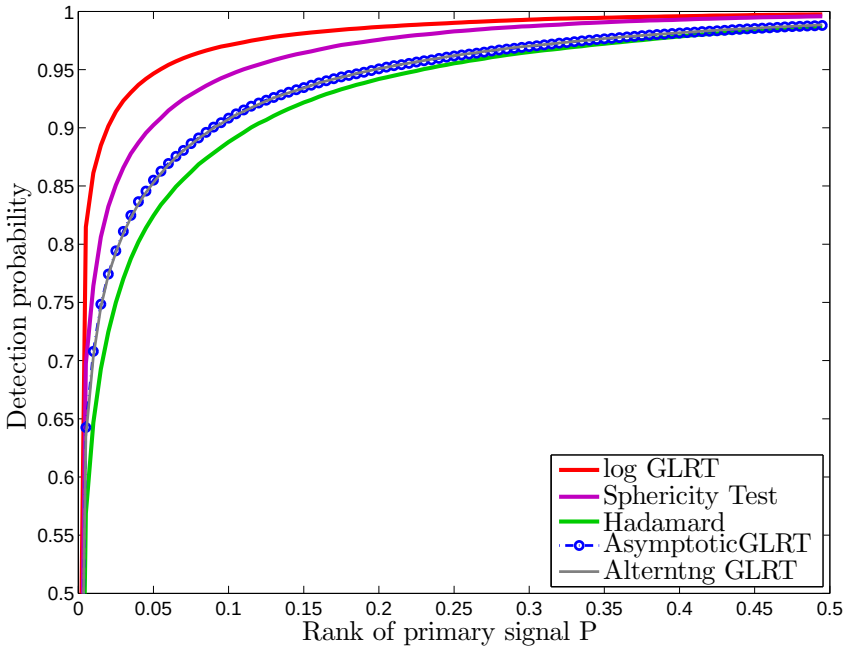

Figure 5.3: ROC for the different detectors (SNR = -8dB, $L = 4$ antennas, $M = 128$ samples, i.i.d noises)

5.1.1.4 Comparison of false alarm probability versus detection probability - non-i.i.d noises

Figure (5.5) compares using the ROC curve the same detectors as the previous figure with a similar scenario, now assuming non-i.i.d noises.

As we can see, the detection performance derived for uncalibrated receivers does not have any difference with respect to the previous figure. This can be explained from the fact that these detectors assume no structure in the noise covariance and therefore, the performance is not affected by the noise distribution.

On the other hand, the detectors designed for uniform noises, i.e. *Sphericity* and *i.i.d-GLRT* suffer a considerable performance degradation since the noise structure does not accomplish with the design assumptions.

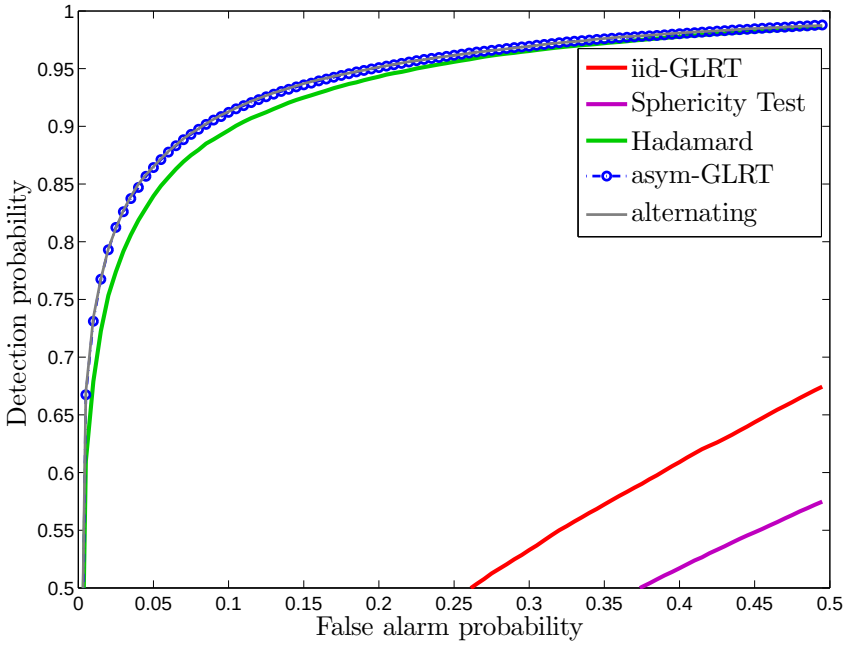

Figure 5.4: ROC for the different detectors (SNR = -8dB, $L = 4$ antennas, $M = 128$ samples, non-i.i.d noises)

5.1.1.5 Comparison of missed detection probability versus SNR

In order to compare the detectors designed with non-i.i.d noises at the receivers, we will represent the missed detection probability for different SNR values. It will be assumed only one transmitter $P = 1$ and $L = 6$ antennas. The result will be represented for two different values of $P_{FA} = 0.01$ and $P_{FA} = 0.1$.

As expected, for very low SNR values the approach designed to the low SNR regime, i.e. *asym-GLRT*, offers as good performance as the *alternating* optimization. As the value of the SNR increases, the *asym-GLRT* does not accomplish with the required low-SNR regime and therefore, the obtained performance is worst than the offered by *alternating* optimization, but the difference between them is very small.

Since the rank $P < L - \sqrt{L}$, the detection performance of the *Hadamard* algorithm is the worst one compared to the previous mentioned detectors.

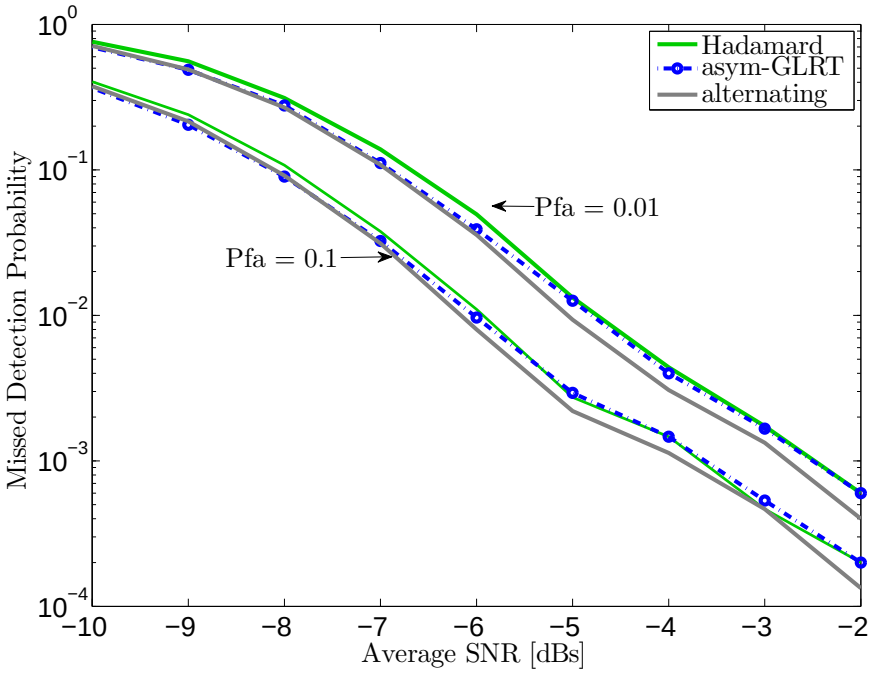

Figure 5.5: Missed detection probability versus SNR for different detectors (SNR = -8dB, $L = 4$ antennas, $M = 128$ samples, non-i.i.d noises)

5.2 Multichannel GLRT detection of stationary processes with arbitrary PSDs

In frequency selective channels, the hypothesis test will be as follows:

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{x}[n] = \mathbf{v}[n], & n = 0, 1, \ldots, N-1, \\ \mathcal{H}_1 &: \mathbf{x}[n] = (\mathbf{H} * \mathbf{s})[n] + \mathbf{v}[n], & n = 0, 1, \ldots, N-1, \end{aligned} \tag{5.10}$$

where $\mathbf{s}[n] \in \mathbb{C}^P$ is the wide sense stationary (WSS) zero mean complex Gaussian primary signal, $\mathbf{H}[n] \in \mathbb{C}^{L \times P}$ is the frequency selective multiple-input multiple-output (MIMO) channel and $\mathbf{v}[n] \in \mathbb{C}^L$ is the WSS zero-mean circular complex Gaussian and spatially uncorrelated additive noise. The SNR for this experiment is defined as follows

$$\text{SNR (dB)} = 10 \log_{10} \frac{\int_{-\pi}^{\pi} \text{tr}(\mathbf{H}(e^{j\theta})\mathbf{H}^H(e^{j\theta})) \frac{d\theta}{2\pi}}{\int_{-\pi}^{\pi} \text{tr}(\mathbf{\Sigma}^2(e^{j\theta})) \frac{d\theta}{2\pi}}. \tag{5.11}$$

In this part we will compare the algorithms that we studied in Chapter 4. As in the previous part, we will compare GLRT detectors under i.i.d. noise assumption. Secondly, GLRT detectors where the noise is assumed non-i.i.d.

The summary of the GLRT algorithms is

- **Freq-log-GLRT**

  GLRT algorithm assuming equal PSDs and rank $P < L - 1$.

$$\begin{aligned}\log \mathscr{L} =& NML \int_{-\pi}^{\pi} \log \left[\frac{\left(\prod_{i=1}^{L} \lambda_i(e^{j\theta}) \right)^{\frac{1}{L}}}{\frac{1}{L} \sum_{i=1}^{L} \lambda_i(e^{j\theta})} \right] \frac{d\theta}{2\pi} \\ & - NM(L-P) \int_{-\pi}^{\pi} \log \left[\frac{\left(\prod_{i=P+1}^{L} \lambda_i(e^{j\theta}) \right)^{\frac{1}{(L-P)}}}{\frac{1}{(L-P)} \sum_{i=P+1}^{L} \lambda_i(e^{j\theta})} \right] \frac{d\theta}{2\pi} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma.\end{aligned} \tag{5.12}$$

- **Freq asym-GLRT**

  GLRT algorithm assuming different PSDs along the antennas, rank $P < L - \sqrt{L}$ and low-SNR regime

$$\log \mathscr{L} \approx NM \int_{-\pi}^{\pi} \sum_{i=1}^{P} \log \beta_i(e^{j\theta}) \frac{d\theta}{2\pi} - NM \int_{-\pi}^{\pi} \sum_{i=1}^{P} \beta_i(e^{j\theta}) \frac{d\theta}{2\pi} + NMP \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \gamma. \tag{5.13}$$

5.2.1 Numerical results

5.2.1.1 Comparison of false alarm probability versus detection probability - equal PSDs

To represent Figure (5.6), let us assume $P = 2$ transmitters and $L = 5$ antennas, which captures $M = 5$ realizations. We will represent two curves for each detector which show the detection performance with different values of SNR and length of the realization N, i.e: 1) *SNR = 2.5dB, N = 20* ; 2)*SNR = -2.5dB, N = 100.*

The figure shows the good detection performance of the *Freq-log-GLRT* algorithm compared to the *i.i.d-GLRT*, which is derived for a frequency flat channel. As known, when the SNR decreases, the detection performance is degraded. This effect can be mitigated by increasing the realization length N but as we can see, it does not have the same effect on the performance of both detectors seen in the figure.

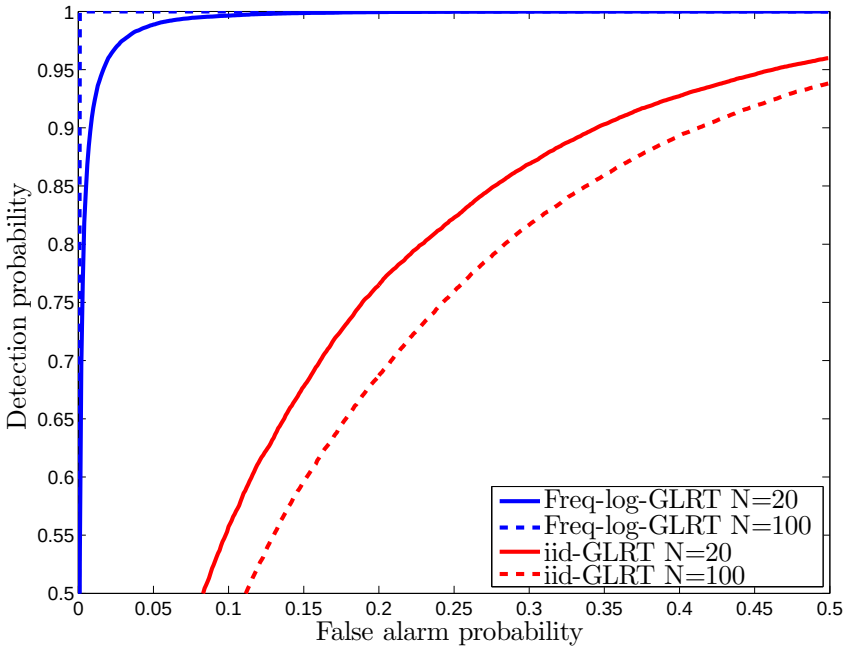

Figure 5.6: Performance comparison of the frequency domain GLRT

5.2.1.2 Comparison of false alarm probability versus detection probability - different PSDs

In Figure (5.7) the scenario is similar to the previous one, now assuming different PSDs along the components, i.e. the parameters are $P = 2$ transmitters and $L = 5$ antennas with $M = 5$ realizations. The represented ROC curves have the parameters: 1) *SNR = 2.5dB, N = 20* ; 2)*SNR = -2.5dB, N = 100.*

As expected, the detection performance of the *asym-freq-GLRT* is better than the *asym-GLRT* for one frequency, as we saw in the previous figure with equal PSDs.

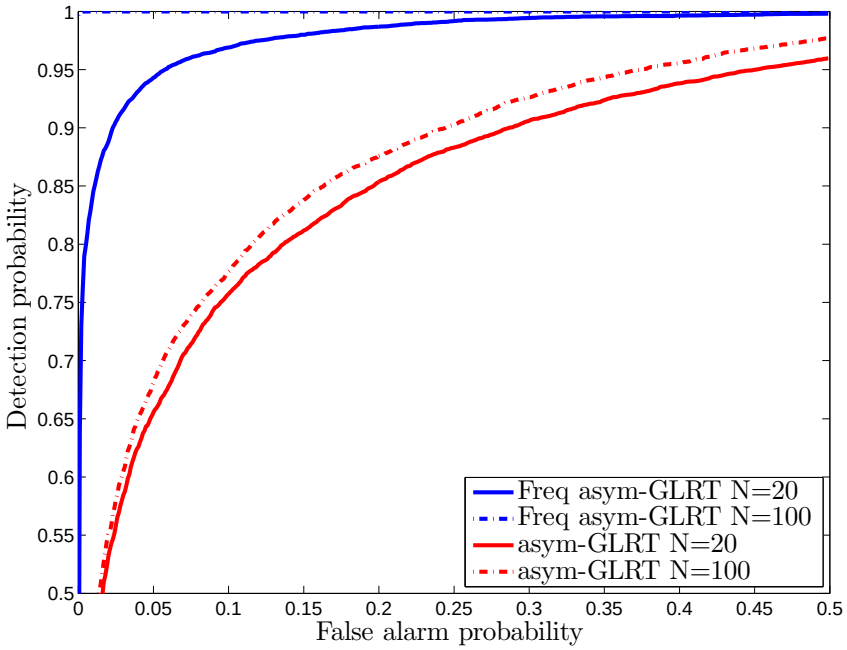

Figure 5.7: Performance comparison of the frequency domain GLRT

Chapter 6

Conclusion

In Chapter 1, we introduced the current problem of spectrum availability. As a solution, we have presented the cognitive radio paradigm, which improves the utilization of the spectrum bands, permitting the access to a secondary user by different techniques. We have briefly introduce three types of cognitive radio: *underlay*, *overlay* and *interweave* cognitive radio, focusing the study on *interweave* in the last chapters of this thesis.

The *interweave* cognitive radio paradigm improves the spectrum utilization allowing the dynamic access to the spectrum bands whenever the primary user is not transmitting. To do that, it is necessary to periodically sense the spectrum using detectors and decide if a secondary transmission is possible. We have considered the spectrum sensing as a binary testing problem, where the presence and absence of signal is represented by two different hypothesis. We have reviewed some detectors as a solution to spectrum sensing in cognitive radio, commenting on their advantages and disadvantages.

Finally, since the generalized likelihood ratio test (GLRT) detector provides good performance, it has been studied in depth in the last chapters of this thesis. We have analyzed the different GLRT algorithms under different assumptions in a multiple-input multiple-output (MIMO) communication system, with P transmitters and L receivers.

In Chapter 3, we have considered a frequency flat channel, i.e. when there is only one given frequency to sense. The GLRT has been reviewed for scenarios, where the noise is independent and identically distributed (i.i.d.) at each of the components. Moreover, we have analyzed scenarios with uncalibrated receivers where the noise is independent and non identically distributed (non-i.i.d). For each scenario, we have obtained an expression to the GLRT detector, whose detection performance has been analyzed in Chapter 5. The most important conclusion in this part is the fact that the detectors designed under non-i.i.d. assumption do not suffer much degradation in the detection performance, even when the noise is i.i.d. On the other hand, the detectors that assume calibrated antennas suffer

a notable degradation in the detection performance when this assumption of i.i.d. noises is not accomplished.

In Chapter 4 the derivation of the GLRT has been studied assuming frequency selective channels, where the temporal structure of the signal is exploited. The problem has been divided into two parts, assuming equal PSDs and different PSDs along the antennas, respectively. As we have seen in the numerical results in Chapter 5, the results given by this frequency GLRT are better than the GLRT, whose design does not take into account the temporal structure of the signal.

Bibliography

[1] Yonghong Zeng. A review on spectrum sensing for cognitive radio: Challenges and solutions. *EURASIP journal on advances in signal processing*, 2009.

[2] Erik G. Larsson Erik Axell, Geert Leus and H. Vincent Poor. Spectrum sensing for cognitive radio. 2012.

[3] Steven M.Kay. *Fundamentals of statistical signal processing. Detection theory.* Prentice Hall, 2013.

[4] A. Goldsmith, S.A. Jafar, I. Maric, and S. Srinivasa. Breaking spectrum gridlock with cognitive radios: An information theoretic perspective. *Proceedings of the IEEE*, 97 (5):894–914, May 2009. ISSN 0018-9219.

[5] Simon Haykin. Cognitive radio: brain-empowered wireless communications. *Selected Areas in Communications, IEEE Journal on*, 23(2):201–220, Feb 2005.

[6] R. Tandra and A. Sahai. Snr walls for signal detection. *Selected Topics in Signal Processing, IEEE Journal of*, 2(1):4–17, Feb 2008. ISSN 1932-4553.

[7] D. Ramirez, G. Vazquez-Vilar, R. Lopez-Valcarce, J. Via, and I. Santamaria. Detection of rank- p signals in cognitive radio networks with uncalibrated multiple antennas. *Signal Processing, IEEE Transactions on*, 59(8):3764–3774, Aug 2011. ISSN 1053-587X.

[8] S. Boyd y L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[9] T. W. Anderson. *Asymptotic theory for principal component analysis*, volume The Annals of Mathematical Statistics, Vol. 34, No. 1 (Mar., 1963), pp. 122-148. Institute of Mathematical Statistics, 1963.

[10] D. Ramirez, J. Via, I. Santamaria, and L.L. Scharf. Detection of spatially correlated gaussian time series. *Signal Processing, IEEE Transactions on*, 58(10):5006–5015, Oct 2010. ISSN 1053-587X.